NOTE TO USERS

Page(s) missing in number only; text follows. Page(s) were scanned as received.

125

This reproduction is the best copy available.

UMI®

Université de Sherbrooke

Dépistage du Cancer de la Prostate: analyse décisionnelle

par

Andriy Moshyk

Département de biochimie

Mémoire présenté à la Faculté de médicine

en vue de l'obtention du grade de

maître és sciences (M.Sc.)

en Sciences Cliniques

Décembre 2004

Your file Votre référence
ISBN: 978-0-494-17339-8
Our file Notre référence
ISBN: 978-0-494-17339-8

Dépistage du cancer de la prostate : une analyse décisionnelle

par Andriy Moshyk (Département de Biochimie Clinique, Université de Sherbrooke)

Introduction

Le cancer le plus répandu et le deuxième plus meurtrier chez les hommes est le cancer de la prostate. Afin d'améliorer les chances de survie des patients, il est nécessaire de faire un dépistage tôt dans la maladie. La stratégie principale de dépistage utilise différents marqueurs qui identifient la maladie chez le patient. Cependant, le choix des marqueurs est très variable. Depuis le début des années 90, moment où une grande évolution s'effectue au niveau des marqueurs, le choix de quels marqueurs sont les plus performants est devenue une tâche fastidieuse. Nous proposons donc une modélisation décisionnelle qui permettra de faire l'évaluation des différentes stratégies et marqueurs existants.

Méthode

Nous avons utilisé la représentation conceptuelle du problème du cancer de la prostate pour faire un modèle en trois phases : dépistage, déterminer le stade de la maladie, traitement. Les données utilisées proviennent d'études systématiques publiées et d'une étude systématique particulière qui vise le dépistage du cancer de la prostate par de nouveaux marqueurs biochimiques. Différentes stratégies alternatives ont été évaluées : l'antigène spécifique de la prostate totale (tASP), ASP complexe (cASP), ASP libre (lASP), le rapport de ASP libre sur ASP totale (l/tASP), le rapport ASP complexe/totale (c/tASP) ainsi que toutes avec/sans touché rectal (TR). Un niveau de sensibilité a été établit à 90% pour tous les tests de dépistage. L'utilité prévisionnelle des stratégies alternatives a été calculée en utilisant la simulation de Monte-Carlo. De

plus, nous avons utilisé le test de Student pour comparer les différentes stratégies de dépistage. Finalement, une analyse de sensibilité avec représentation en diagramme de tornade a été appliquée à la survie des patients en ce qui concerne les caractéristiques de la population. Deux logiciels pour la construction du modèle de décision (ReasonEdge et Data 3.5) ont été utilisés.

Résultat

Une approche d'intégration des évidences a été utilisée pour joindre les différentes parties du modèle et l'information probabiliste des sources hétérogènes. Le modèle a été simulé pour estimer le coût d'un programme de dépistage annuel de 5 ans pour les scénarios suivants (moyenne, écart type): tASP+TR - ($641, $372), cASP+TR ($630, $360), tASP seulement - ($545, $318), cASP seulement - ($535, $302), tASP+TR+c/tASP - ($652, $375), tASP+TR+l/tASP - ($655, $379). Une différence significative entre les programmes de dépistage avec TR et sans TR a été détectée ($p<0,05$). Aucune différence significative entre ASP totale et ASP complexe dans les stratégies semblables (ASP totale vs. ASP complexe avec TR, ASP totale vs. ASP complexe sans TR) n'a été détectée. L'utilisation de l'analyse de sensibilité avec la représentation en diagramme de Tornade a prouvé que la stabilité de la conclusion concernant la survie globale des patients atteints du cancer de la prostate dépend principalement de deux facteurs: la probabilité annuelle de décès pour les groupes suivant les traitements T1/T2 et M1 et la probabilité des métastases distantes.

Conclusion

Différentes méthodologies (modélisation décisionnelle et revue systématique) ont été examinées pour l'évaluation du dépistage du cancer de la prostate. Le processus de modélisation a été basé sur la création du modèle conceptuel du problème et le choix d'informations probabilistes basées sur la relation structurale entre les éléments du modèle de décision. Des lignes directrices de représentation ont été utilisées afin d'éviter les problèmes de transparence et d'augmenter la réutilisation du modèle. De plus, le modèle résultant est généralisable car il est possible de lui poser différentes questions. Finalement, les stratégies de dépistage et l'examen des facteurs importants pour les décisions ont été évaluées. L'examen des influences du dépistage sur la détection du stade du cancer aidera l'estimation de l'impact de ce dépistage sur la survie de la population.

Members of Jury

1. Dr. Andrew Grant, MD, PhD	Department of Biochemistry (Faculty of Medicine, Université de Sherbrooke)
2. Prof. Marie-France Dubois, PhD	Department of Community Health Sciences (Faculty of Medicine, Université de Sherbrooke)
3. Prof. Casimir A. Kulikowski, PhD	Department of Biomedical Engineering, Rutgers, The State University of New Jersey 110 Frelinghuysen Road, Piscataway, NJ 08854 8019, USA

Content

List of Figures

List of Tables

List of Abbreviations

ACS-NPCDP – The American Cancer Society – National Prostate Cancer Detection Project

BMJ – British Medical Journal

CHUS – Centre Hospitalier de l'Université de Sherbrooke

cPSA (or cASP) – complexed prostate specific antigen (or l'antigène spécifique de la prostate complexe)

c/tPSA (or c/t ASP)– complexed-to-total prostate specific antigen ratio (or le rapport de ASP complexe sur ASP totale)

CT - computerized tomography

DRE (or TR) – digital rectal examination (or touché rectal)

ERSPC - European Randomized study of Screening for Prostate Cancer

FN – false negative

FP – false positive

fPSA (or lASP)– free prostate specific antigen (or l'antigène spécifique de la prostate libre)

f/tPSA (or l/t ASP) – free-to-total prostate specific antigen ratio (or le rapport de ASP libre sur ASP totale)

GS – Gleason Score

MRI - magnetic resonance imaging

PAP – prostatic acid phosphatase

PC – prostate cancer

PLCO - Prostate, Lung, Colorectal and Ovarian Cancer Screening Trial

PSA (or ASP) – prostate specific antigen

QALE – quality adjusted life expectancy

QALY – quality adjusted life year

RCT – randomized clinical trial

ROC – receiver operating characteristics

TN – true negative

TNM - American Joint Committee on Cancer's morphological classification

TP – true positive

tPSA (or tASP) – total prostate specific antigen (or l'antigène spécifique de la prostate totale)

y.o. – years old

I. Introduction

The subject of this master's thesis is the development of a decision model that concerns the screening, staging and treatment phases of prostate cancer. This chapter starts with consideration of major evidence on the importance and characteristics of prostate cancer. This is followed by a discussion on decision modeling in health care. The modeling domain and clinical domain sections contain core information about the problem perspectives studied in relation to the developed model. The subsequent section provides different points of view on the use of information and evidence in health care decision making. The analysis of prostate cancer decision models published to date finishes the chapter.

Importance of prostate cancer

Prostate cancer is a growing health problem with considerable economic consequences (VARENHORST et al., 1994). Prostate cancer is now the most common cancer and the second most common cause of death from cancer among men (NEHEMAN et al., 2001, RECKER and LUMMEN, 2000, WINGO et al., 1995). Radical treatment is usually possible for organ confined disease (NEHEMAN et al., 2001, RECKER and LUMMEN, 2000).

Han et al. (2001) have shown that a screening program for prostate cancer can improve diagnosis of early stages. Organ confined cancer is more frequent for cases identified by the screening program. They demonstrated a biochemical recurrence-free survival advantage due to an improved therapeutic outcome and lead time bias (HAN et al., 2001).

Populational impact of screening

Several publications exist in the research literature to investigate prostate cancer problem and the impact of screening, diagnosis and treatment. The studies differ by population and study design. The introduction of the prostate specific antigen screening test (PSA) and its impact on the subsequent management of disease has been increasingly studied. The impact of screening remains a centre of attention.

Sarma and Schottenfeld (2002) conducted a retrospective study linking demographic data from a US population with an implementation of prostate cancer screening in clinical practice. Prostate cancer incidence increased steadily from 1981 to 1989, with a steep increase in the early 1990s, followed by a decline. The exaggerated rate of increase in the early 1990s in prostate cancer incidence was transient and likely a result of increased detection of preclinical disease that was prevalent in the general population (SARMA and SCHOTTENFELD, 2002).

Results of two studies on the Quebec population were published in last five years. The conclusions derived from them illustrate the controversy about screening. Perron et al. (2002) did a retrospective study on birth cohorts of the Quebec province using regression modeling on relative mortality with factors including an exposure to prostate cancer screening. According to them, the difference in prostate cancer mortality is not attributable to total PSA (tPSA) screening. If tPSA screening is effective in preventing or postponing death from prostate cancer, its impact at a population level has yet to be felt. They suggested that there may be other explanations for the recent decline in prostate cancer mortality, consisting primarily of changes in disease management and in hormonal treatment of advanced disease (PERRON et al., 2002).

The second study, also from the Quebec population, shows another point of view. If tPSA screening is started at the age of 50 years, annual or biannual tPSA alone is highly efficient to identify the men who are at high risk of having prostate cancer. This prospective study was conducted on a population of men randomly allocated into screening and non-screening groups with ratio 2:1. Patients were randomly selected from the electoral list and invited by mail without any public announcement. Labrie et al. (1999) demonstrated that early diagnosis and treatment permits a decrease in deaths from prostate cancer (LABRIE et al., 1999).

Vis (2002) summarized the study design critique of these two trials. In his view the reported decline may be the result of increased use of curative treatment before the implementing of tPSA screening and the availability of new treatment options for advanced prostate cancer. Changes in diet, lifestyle and environmental conditions, and the incorrect labeling of deaths from other causes could be also attributable (VIS, 2002). The scientific community is also waiting for the results of two ongoing randomized controlled trials in North America and Europe (European Study of Screening for Prostate Cancer (ERSPC)[1] and Prostate, Lung, Colorectal and Ovarian Cancer Screening Trial (PLCO)[2]). These studies are supposed to provide the definitive evidence whether tPSA based screening is beneficial for patient survival or not. Results from these studies are expected to be available by 2006-2008.

Diagnostic problems in prostate cancer

There are other questions yet to be discussed which are relevant to the current screening programs. A prostate biopsy is an obligatory test to confirm the presence of prostate cancer. Screening programs are aimed to select appropriate groups of patients for

---

[1] See http://www.erspc.org/ . Accessed on the 19th of July 2004.
[2] See http://www3.cancer.gov/prevention/plco . Accessed on the 19th of July 2004.

prostate biopsy. Screening programs have low specificity, which cause 65-75 % negative biopsies for some groups of patients (e.g. 4-10 mg/l range of prostate specific antigen, or tPSA) (POTTER et al., 2001). Roberts et al. (2000) supported the use of diagnostic techniques in order to reduce the number of negative biopsies and improve cancer yield in younger men. Repeated negative biopsies assumed as unnecessary might frustrate a patient if the screening program is provided on serial basis (e.g. annually). It also has an influence on increase of health care costs (ROBERTS et al., 2000). This could give an explanation why serial screening programs are not commonly adopted.

Hence, despite the fact that prostate cancer is well represented in the scientific literature and is a problem with large impact on population health, grey zones are still left. Major topics during the last 10 years after bringing tPSA screening into practice, were screening effectiveness, optimal use of screening markers and evaluation of measures of free or complexed PSA derivatives vs. total PSA.

Prostate cancer domain

This section provides a review of existing evidence for the prostate cancer domain, which is necessary for understanding the structural assumptions for the prostate cancer models.

Natural history of prostate cancer

In general it is a slow growing cancer (KESSLER and ALBERTSEN, 2003). Age is the most important factor associated with the cancer development (PORTER et al., 2002). For men between 40-49 years old a histological prevalence of the prostate cancer is near 12%, but for the men after 80 years old it is up to 43% (COLEY, 1997a). Near 95% of prostate cancers are adenocarcinomas. Most prostate cancer (75%) arises in the peripheral

zone of the prostate gland, nearly 15% develops in the transition zone and the remainder arises in the central zone (AUGUSTIN et al., 2003). Despite the fact that prostate cancer cells have been detected in almost 1/3 of men over 50, in many cases disease does not reach clinical stage (SCARDINO et al., 1989). A candidate for predicting risk of developing cancer is prostatic intraepithelial neoplasia (DEMARZO et al., 2003).

The prostatic capsule acts as an initial barrier to local invasion of the surrounding tissues. If localized within the prostate capsule, the cancer is assumed eligible for radical treatment using prostatectomy or radiotherapy, which is associated with a good prognosis. Once the capsule is invaded, the disease is viewed as locally advanced prostate cancer. Invasion of vascular and lymphatic tissue introduces the chance of metastatic spread of disease. Lymph from the prostate gland drains into lymph nodes in the pelvis, groin and lower back and these lymph nodes become common sites for metastasis. Secondary disease from prostate metastases mainly arises in the bones (FRYDENBERG, 1997).

Screening and diagnosis

Prostate cancer is usually described as induration of the prostate on digital rectal examination (DRE) if palpable (PRESTI, Jr. et al., 2000). The implementation of serum testing for tPSA has significantly improved the ability to detect cancer. tPSA used as pre-screening followed by DRE is highly efficient in detecting prostate cancer at a localized stage (CANDAS et al., 2000). While serum tPSA testing combined with DRE has good sensitivity for detecting prostate cancer, specificity is low due to the non-cancer specific elevation of tPSA, which is attributable to benign prostate disease (BRAWER, 2000).

Hoedemaeker et al. (2000) suggested that screening for prostate cancer leads to an increase in surgical treatment for relatively small tumors that have a higher probability of being organ confined. The frequency of positive lymph nodes at operation decreases

dramatically and the proportion of organ confined tumors after surgery increases, there is a shift from tumors with Gleason Score (GS) 8-10 towards lower grade tumors at radical prostatectomy (HOEDEMAEKER et al., 2000).

According to Candas et al.'s (2000) results from a cohort study of 11,811 participants, there is a 7-fold decrease in prevalence of prostate cancer at follow-up visits done up to 11 years. tPSA alone allowed to find 90.5% and 90% of cancers at first and follow-up visits, respectively, compared to 41.1% and 25% by DRE alone. This means that tPSA is not losing performance due to eliminating cancer cases from the follow-up population (CANDAS et al., 2000). Rietbergen et al. (1998) also indicate that the chance of diagnosing prostate cancer in men by a positive DRE is decreased at follow-up visits in comparison to the first visit for serial screening program (RIETBERGEN et al., 1998).

PSA and derivates as screening markers

tPSA consists of 3 forms: free PSA, PSA complexed with alpha 1-antichymotrypsin and PSA complexed with beta 2-macroglobulin (CHRISTENSSON et al., 1993). The beta 2-macroglobulin bound form cannot be detected by monoclonal and polyclonal antibodies prepared against PSA. Of the 3 major serum forms of PSA, only free PSA and PSA complexed with alpha 1-antichymotrypsin are immunodetectable by current commercial assays (CATALONA et al., 1995, OESTERLING, 1995).

In the commonly accepted diagnostic zone of 4 to 10 µg/l total PSA, prostate cancer is present in 25% of patients. To maintain acceptable sensitivity a high number of biopsies are being performed. This tPSA range is often called a grey zone. Such patients are viewed by many researchers as a potential population where specificity of screening program could be improved (BRAWER et al., 2000). Several modifications to screening programs have been recently suggested like age-adjusted tPSA cut-offs, tPSA density (LENTINI et al.,

1997, PIZZOCCARO et al., 1994), tPSA adjusted for volume of transition zone (GUSTAFSSON et al., 1998, KIKUCHI et al., 2000, LUBOLDT et al., 2000, MAEDA et al., 1997). However none of these proposed approaches has gained common practice use (FLESHNER et al., 2000).

Several studies have demonstrated that the proportion of free or complexed-to-total PSA enhances the clinical usefulness of PSA testing for the early detection of prostate cancer, and it may reduce unnecessary biopsies (MILLER et al., 2001, MITCHELL et al., 2001, PRESTIGIACOMO and STAMEY, 1997, TANGUAY et al., 2002, VASHI and OESTERLING, 1997, WOODRUM, 1998). Complex-to-total PSA (c/tPSA) and free-to-total PSA (f/tPSA) ratio were found similar in performance (LEIN et al., 2001, OKEGAWA et al., 2000). Both of them can significantly improve detection of prostate cancer especially in the 4-10 mg/l tPSA range.

The normal reference range of free-to-total PSA ratio reported by Catalona et al.(1995) and Oesterling (1995) was 23 to 31 percent (CATALONA et al., 1995, OESTERLING, 1995). Percent of free PSA may increase the specificity of tPSA testing without sacrificing the cancer detection rate (HIGASHIHARA et al., 1996, WOODRUM, 1998). Brawer et al. (2000) further demonstrated that the complexed PSA method as a single measurement enhances specificity for detecting prostate cancer comparable to the measurement of percent free PSA. These findings suggest that complexed PSA may serve as a single assay replacement of the measurement of total PSA (BRAWER et al., 2000). Summarizing the presented evidence the order of biochemical markers as they were introduced into practice has been the following [PSA (or PSA density)→ fPSA(or free-to-total ratio) → cPSA (or complexed-to-total ratio)].

Biopsy

Djavan et al. (2001) in a prospective study on prostate cancer detection with repeated biopsies for men with total PSA between 4 and 10 μg/l found a 10% cancer rate on second biopsy 6 weeks after a first negative biopsy. The initial cancer rate on the first biopsy was 22% (DJAVAN et al., 2001). These findings suggest that needle prostate biopsy is an imperfect test for determination of prostate cancer for a screened population. However this test is still the best available for the cancer detection. Currently second biopsy just after the first one to find missed cancers is not judged necessary. Some missed cancers (up to 10 - 11%) might be diagnosed next year during a next round of the screening.

Trans-rectal ultrasound of prostate

Trans-rectal ultrasound of prostate (TRUS) is not warranted in men with normal DRE and tPSA results (BABAIAN et al., 1993). Various studies have suggested to avoid TRUS as the first order test (e.g. when selection of general population for prostate biopsy was done using three tests such as tPSA, DRE and TRUS independently) (HIGASHIHARA et al., 1996).

Serial prostate cancer screening

Most of the published studies show results on 1 year (or single measurement) screening. At any given visit, the tPSA levels of approximately 25% of men with initially elevated levels had decreased to less than 4.0 μg/l. Of all prostate cancer detected, 85% were detected during the first 2 years of screening. After 3 to 4 years of screening, the proportion of men with abnormal test results decreases substantially, the cancer detection rate decreases even more to approximate the expected prostate cancer incidence rate. There is a shift to detection of earlier-stage disease (SMITH et al., 1996).

Rietbergen et al. (1999) provides comparative evidence on the impact of prostate cancer screening. Comparison of the characteristics of prostate cancer between two populations (screening general population vs. population without screening) revealed reduction in advanced stage disease primarily due to the number of metastatic cases. Authors suggested further evaluation of stage reduction and disease specific mortality (RIETBERGEN et al., 1999).

Staging

There are two main stage classifications of prostate cancer (TNM[3] and Whitmore). TNM is a way of describing the size, location and spread of a tumor. T denote the primary tumor according to its size and location. N refers to whether the cancer has spread to the lymph nodes that drain fluid from that area. M represents whether there are metastases in distant areas (e.g. M1 cancers) (STAMEY et al., 1998). Cancer development time consists of two stages "latent" and "clinical" cancer. Another classification was suggested by Whitmore. The correspondence between these two classifications is presented in Table 1.

[3] American Joint Committee on Cancer's TNM classification

Table 1. Current classifications of prostate cancer (FRYDENBERG, 1997, HAN et al., 2000, O'DOWD et al., 1997).

Whitmore	AJCC/TNM	Characteristics of tumor
A	T1	Clinically not palpable or visible by imaging tumor
	T1a	Tumor incidental, found in 5% or less resected tissue
	T1b	Tumor incidental, found in more than 5% resected tissue
	T1c	Tumor identified by needle biopsy (because of high tPSA)
B	T2	Tumor confined within the prostate gland
	T2a	Tumor involves one lobe
	T2b	Tumor involves both lobes
C	T3	Tumor extends through the prostatic capsule
	T3a	Extracapsular extensions (unilateral or bilateral)
	T3b	Tumor invades seminal vesicles
D	T4	Tumor is fixed or invades adjacent structures

There is a continuous discussion on what should be viewed as latent and clinical cancers. In general the impact of presence of the cancer on the expected life length is the only parameter for assessing clinical significance. There is an increasing probability to die from other reasons with increasing age rather than to die from prostate cancer.

Detailed information about prostate cancer profile (stage distribution, age stratification, mortality) is very useful for an estimation of general impact on population health after tPSA implementation. Amling et al. (1998) have examined a large population (5,568 referred patients with prostate cancer, who underwent pelvic lymph nodectomy and radical retropubic prostatectomy between 1987 - 1995) who had adenocarcinoma of the prostate. The percentage of patients with stage T1c prostate cancer (this stage can be detected by tPSA only) increased, and stage T3 cancer decreased. At the same time histological grade decreased and the proportion of pathological organ-confined disease increased, which is similar to clinical stage changes. Five-year progression-free survival was 85% and 76% for patients with clinical stage T1c and T2, respectively. Radical prostatectomy experience has shown a significant migration to lower-stage, more differentiated, more often organ-confined

prostate cancer at the time of initial assessment after tPSA testing has appeared in clinical practice. Cancer-free survival associated with tPSA-detected cancer (T1c) is superior to that with palpable tumors (T2). Due to study design limitations, Amling et al. (1998) also suggested that improved long-term cancer-specific survival remains to be confirmed with longer follow-up (AMLING et al., 1998).

Treatment selection is influenced by local stage assessment. Most of the time, clinicians must distinguish between pathologically (p) confirmed organ-confined disease (pT1-2) and non-organ-confined disease (pT3-4) (PRESTI, Jr., 2000). Patients with organ-confined disease can be treated with surgery or radiation therapy, patients with extra capsular extension or seminal vesicle invasion are not surgery candidates. They can be treated with radiation therapy, hormonal therapy or a combination of both (YU and HRICAK, 2000).

Understaging may result in ineffective local treatment (surgery or radiation therapy) with the unnecessary risks and costs. Overstaging may result in withholding potentially radical therapy when a tumor might be amenable to definitive local treatment (KINDRICK et al., 1998).

Clinical stage

T stage is the clinical determination of local extension of disease primarily by digital rectal examination. The most widely used clinical stage classification system for prostate cancer was introduced by Whitmore. The clinical T stage only indirectly helps the urologist make important pre-treatment diagnostics decisions. DRE lacks specificity in the determination of organ confined (or sensitivity for non-organ confined) disease. But the predicted clinical stage correlates with pathological stage (O'DOWD et al., 1997).

A prostate cancer preoperative stage underestimates the final pathology stage in approximately 40-50% of the cases (RUBIN et al., 1997). The frequency of pathologic

understaging is partly related to a clinical stage ranging from 30% in clinical stage T1b to 60% in clinical stage T2 disease. Non-palpable (T1c) prostate cancer is the most commonly diagnosed stage of disease at presentation today because of the widespread use of tPSA (PRESTI, Jr., 2000). Clinical stage is effective for identifying advanced disease. (YU and HRICAK, 2000).

Tumor grade (Gleason score)

The Gleason grading system is the most commonly used grading system for prostate cancer histology in North America. The pathologist assigns a primary grade to the pattern of cancer that is most commonly observed and a secondary grade to the pattern of cancer that is the second most commonly observed in the specimen. Grades range from 1 to 5. The Gleason score is obtained by adding the primary and secondary grades together. Well-differentiated tumors have a Gleason sum of 2 to 4, moderately differentiated tumors have a Gleason score of 5 to 6, whereas poorly differentiated tumors have Gleason score of 8 to 10. The likelihood of having organ-confined disease decreases with increasing tumor grade (PRESTI, Jr., 2000).

PSA for preoperative stage determination and assessment of imaging needs

tPSA does not have perfect predictive capacity for a particular clinical stage. However total serum PSA correlates directly with advancing clinical and pathological stage of prostate cancer (PARTIN et al., 1993). Imaging is quite an expensive procedure. tPSA has been used to identify the group of patients where imaging would be more efficient. Men with tPSA level less than 4 μg/l generally have organ-confined disease, whereas approximately 50% of patients with tPSA levels over 10 μg/l have extra-capsular extension. (PRESTI, Jr., 2000). According to Morote et al.'s (1997) study, tPSA can be successfully used to eliminate

the radionuclide bone scan in 40 % of patients with newly diagnosed prostate cancer (MOROTE et al., 1997).

Other methods for preoperative stage determination

Combination of significant clinical information was used by Partin et al. (1993) to create nomograms for prediction of pathological state and justify the imaging needs for prostate cancer patients. The purpose was to improve outcome prediction. Such predictive models can be used as a diagnostic test itself. It is a very "easy to use" method. Using a logistic regression modeling approach, Partin et al. (1993) demonstrated that total serum PSA, when combined with Gleason grade and initial clinical stage assessed during digital rectal examination (DRE), provided the best separation among pathological stages (e.g. capsule penetration, extra-prostatic spread, metastases at lymph nodes) compared to any univariate independent variable (PARTIN et al., 1993).

Imaging

The imaging studies are used to identify metastases and/or extra-capsular extension in men with newly diagnosed prostate cancer. This allows identification of patients for whom the definitive treatment will not provide additional survival advantages. Wide variations exist in the use of Gleason Score and serum tPSA in imaging studies. Physicians performed radionuclide bone scans, computerized tomography (CT) and magnetic resonance imaging (MRI) on many men with newly diagnosed prostate cancer as part of the initial stage evaluation to determine whether disease extends beyond the prostate capsule to pelvic lymph nodes or bone (ALBERTSEN, 2000).

Bone scan and computer tomography

Traditionally the radionuclide bone scan has been the cornerstone of prostate cancer stage determination. Previous (before the "PSA era") widespread use of bone-scan imaging was certainly reasonable, even in asymptomatic patients (LEE and OESTERLING, 1997).

Although the risk of a positive bone scan increased with increasing state and grade, tumor stage and grade were poor predictors of positive bone scan according to results of the Gleave et al. (1996) study. Up to 4% of patients with clinically confined or well-differentiated to moderately differentiated tumors had positive scans (GLEAVE et al., 1996).

Several studies on bone scan and CT suggest that these should only be ordered for men with newly diagnosed prostate cancer with tPSA greater than 20 μg/l or tPSA greater than 10 μg/l and Gleason scores 8 to 10. Such populations with higher risk of extra-capsular extension have positive yields greater than 10 % on bone scans (ALBERTSEN, 2000, LETAIEF et al., 2000). Several authors stated that pelvic CT and bone scans for the stage determination are not advocated for the patients with a tPSA level of less 20 μg/l (LEVRAN et al., 1995, LORENTE et al., 1999, STOKKEL et al., 1998). At the same time other authors suggested decreasing tPSA cut-off till 10 μg/l for asymptomatic, newly diagnosed patients (ATAUS et al., 1999, LEE and OESTERLING, 1997, O'DOWD et al., 1997).

Cross-sectional imaging with computed tomography has not proven to be very sensitive for evaluation of extra-capsular disease extension. The very low predicted rate of seminal vesicle invasion and lymph node metastasis determined by the combination of Gleason's score, clinical stage, and tPSA suggests that little benefit is obtained from cross-sectional imaging in patients with rather well-differentiated lesions and tPSA less than 10 μg/l (MANYAK and JAVITT, 1998). Yu and Hricak (2000) suggested that there is no

consensus and there is not enough evidence for guidelines for evaluation of local prostate cancer extent imaging (YU and HRICAK, 2000).

Pelvic lymph node dissection

Dissection of pelvic lymph nodes is the time-proven method for assessment of node involvement. It usually accompanies radical prostatectomy, and can be done independently (by laparoscopy or by mini-laparotomy) before radiotherapy, perineal prostatectomy or if the tPSA is more than 20 μg/l or if the cancer is poorly differentiated or there is clinically advanced local disease, or both. The incidence of disease metastatic to lymph nodes correlates directly with clinical stage. Patients with a tPSA of 20 μg/l or more, Gleason score 8 or more and abnormal digital rectal examination have a high risk for lymph node metastases (PARTIN et al., 1993, WOLF et al., 1993b).

Pathological stage

The pathological examination of radical prostatectomies and pelvic lymph node specimens provides the most accurate description of the extent of disease available (O'DOWD et al., 1997). This can suggest treatment modification in order to improve patient survival if cancer was understaged. According to some recommendations, preoperative lymph node examination can suggest that operation should be stopped before actual prostatectomy.

Treatment

The most important decision to be made in the patient with newly diagnosed prostate cancer is whether or not definitive treatment is necessary. Patients with clinical stage T3 disease are not generally considered surgical candidates, whereas patients with clinically localized disease are considered potential surgical candidates (YU and HRICAK, 2000).

Patients with prostate cancer localized to the pelvis without nodal or distant metastases can be treated with radiation therapy (FRYDENBERG, 1997).

Hormonal therapy is the mainstay of treatment when the patient has lymph-node metastases or disseminated metastases. Bilateral orchidectomy has been the standard treatment for testosterone reduction. Analogues of luteinising-hormone releasing hormone are commonly used for chemical castration and have equivalent effect to bilateral orchidectomy. A combination of medical or surgical castration and anti-androgen therapy can be used to block both testicular and adrenal androgen activity (FRYDENBERG, 1997).

Practice evaluation and decision making

These days clinicians are often guided by peer reviewed guidelines provided by trusted health institutions or professional boards. Guideline content is established by evaluation of best clinical practice results based on published studies in health care. Guidelines review diagnostic tests, treatment and management of disease options.

Common sense suggests that diagnostic technologies should be disseminated only if they are less expensive, produce fewer untoward effects and are at least as accurate as existing methods. They should eliminate the need for other investigations without loss of accuracy. Providing results, acceptable by clinical community, ideally requires a randomized controlled trial in which patients receive the new test or an alternative diagnostic strategy (GUYATT et al., 1986). Multiple evidence on the same subject is often accepted by the research community. However heterogeneity in study design can make it difficult to combine data from different studies.

Evaluation of diagnostic tests

Galen (1982) has suggested four levels for evaluation of laboratory tests. The first level is analytical evaluation of the laboratory test, followed then by diagnostic analysis with evaluation of sensitivity, specificity and receiver operating characteristics (ROC) curve. The third level is operational analysis with evaluation of outcome of positive and negative results and efficiency. The last level contains medical decision making analysis with evaluation of threshold probability, cost-benefit analysis and decision analysis modeling (GALEN, 1982).

Measuring diagnostic effectiveness (the second level in Galen's classification of levels from above) has become routine for implementation of every new test in clinical practice. Sensitivity is the ability of a test to correctly identify individuals who have a given disease or disorder. This is translated to a formula as: TP/(TP+FN), where TP is for true positive, FN is for false negative. Specificity is the ability of a test to correctly exclude individuals who do not have a given disease or disorder. This is translated into a formula as TN/(TN+FP), where TN is for true negative, FP is for false positive. The area under the ROC curve is defined by a curve of sensitivity and 1-specificity at various threshold values. ROC analysis is the dominant technique for evaluating the suitability of diagnostic techniques for real applications (METZ, 1978).

McIntosh et al. (2002) highlighted problems associated with translating a potential screening biomarker from the laboratory to its use in patient care. Such application may require an algorithm or screening rule or even protocol for its application. Any practical screening algorithm must do so with strict controls on test specificity to avoid false-positive results, and unnecessary patient alarm and risk. The author also indicated the importance of longitudinal screening programs. Such programs, where prior tumor marker values and

trends are analyzed, improve the diagnostic performance over a single determination (MCINTOSH et al., 2002).

Even using a marker that can distinguish patients eligible or not eligible for specific intervention may not guarantee survival difference. Estimates of survival difference can be replaced by a surrogate end-point if the size and study period relevant to survival end-point is not realistic (SARGENT and ALLEGRA, 2002).

Study design for survival comparison has a major impact on study validity. To evaluate the hypothesis that a new diagnostic test or strategy is beneficial, randomized controlled trials with two arms are recommended (GUYATT et al., 1986). Running a RCT usually requires a lot of resources. Measuring survival advantage requires long follow-up studies. Repeat studies for every new diagnostic marker or therapeutic changes risk to be wasteful.

Benoit and Naslund (1997) suggested if men aged 50 to 70 years potentially benefit the most from tPSA screening this benefit would not be realized until these men are in their seventh and eighth decades of life (BENOIT and NASLUND, 1997). This study underlines the timeline aspect for studies on evaluation of cancer markers.

There are various heterogeneous approaches at all phases of prostate cancer management (from screening through pretreatment diagnostics to treatment outcome). No single clinical study is able to evaluate the various approaches for diagnostics (screening, staging) at the same time due to study population size limitations. There are also continuous needs for evaluation of new approaches. If a study has been conducted before, repetition seems to be not appropriate from a research point of view (e.g. "reinventing the wheel", no innovation). Such repetitions are also not popular for economic reasons.

Modeling has been proposed as a less expensive way to evaluate different strategies before conducting real clinical studies, because using a modeling approach, the answer to research questions is based on integration of data collected from a number of well undertaken studies (KWOK et al., 2001, SIMPSON, 1994).

Decision modeling in health care

Graphs are a natural way for information representation. Graphical modeling for decision support is getting more widespread. Among all quantitative decision making methods, this approach is a way to think of and communicate on the underlying structure of the domain in question. It also helps the researchers to focus on structure rather than calculations (SMYTH, 1997).

The structure of a graphical model clarifies the conditional independencies in the implied probability models, facilitating model assessment and revision (SMYTH, 1997). Presence of disease or any other condition in any step of a diagnostic protocol can be defined as a probability (value or distribution) (PAUKER and KASSIRER, 1980).

An ideal decision-analysis model includes all important available interventions and defines and discloses the analyst's time frame and outcome assessment perspective. The results can be monitored and, if necessary, adjustments can be made after completing the model if new evidence is available or due to other reasons (NIELSEN and JENSEN, 1999).

The basis for decision modeling is expected utility theory. Expected utility theory is suggested to evaluate choice of different patient oriented medical strategies (CLAXTON et al., 2001). This theory states that we should always choose an alternative that maximizes the expected utility (NIELSEN and JENSEN, 2000). According to Claxton et al. (2001) the choice of strategy and decision for clinical study design and practice evaluation should be based on expected utility. The authors have applied this theory for estimating the needs for

conducting additional research or acquiring additional information through assessing uncertainty surrounding outcomes of interest (CLAXTON et al., 2001).

Decision analysis is best applied in decisions when other factors in addition to acquisition costs are important in determining overall intervention or diagnostics costs for several alternatives. Buxton et al. (1997) listed situations were modeling can be useful. They include a variety of data sources (evidence from trials, systematic reviews of trials) and possibility of application to different clinical settings (BUXTON et al., 1997).

So the decision maker should start the modeling project from the identification of alternative decisions and all factors which could influence the decision. Then probabilistic and outcome data should be identified according to the criteria suggested by the decision maker.

Adaptation of existing published models can be a good approach for decision-modelers. Fewer resources are necessary for this approach, since the underlying structure and variables have already been established. This approach focuses specifically on tailoring an existing model to meet specific needs. But models are often difficult to reproduce because they are not thoroughly described in the literature (an issue of transparency) (SANCHEZ and LEE, 2000). To be useful, prostate cancer treatment models must be based on acceptable structural assumptions, contain valid data and be understandable to clinical experts (SIMPSON, 1994).

Decision analysis modeling is an economically attractive method for practice evaluation that uses different sources of evidence to draw the conclusions about the new technologies or approaches used in health care.

Modeling principles

Decision trees and influence diagrams are two approaches for graphical modeling of decision problems. They offer different perspectives on the same problem using the same mathematical relationship. Each represents certain dimensions explicitly and can sometimes hide other dimensions from view. Influence diagrams detail well the relations among many parameters; decision trees show sequential paths and their branching. Using the two views at the same time is helpful for understanding a decision model (HELFAND and PAUKER, 1997).

Elements of the decision model

The following elements are considered as necessary elements for a decision model: decision node, chance node and utility node. The decision (one or many in the same decision model) is usually represented as a rectangular node. The decision nodes correspond to decision variables and represent alternative actions under the direct control of the decision maker. In the decision tree, the arcs leaving the decision node indicate the possible decisions available at this decision node. In the influence diagram, the alternative scenarios are described at the decision node, but not represented as arcs leaving the node.

The chance nodes (drawn as circles) correspond to chance variables, and represent events which are not under the direct control of the decision maker. Each chance node has outcomes associated with such an event. For the decision tree, the arcs leaving the chance nodes represent outcomes for every particular chance node. The numbers on the arcs leaving chance nodes are the probabilities of the outcomes to appear. In influence diagrams, the arcs connecting chance nodes represent conditional dependency between them. The absence of

the link between chance nodes means that there is conditional independence between events. The outcomes and probabilities for these outcomes to appear are hidden in underlying tables for every chance node. A conditionally independent chance node has a 1*n table where *n* is the number of outcome for a particular chance node. Multiple links from other chance nodes require hierarchical tables, e.g. a chance node with *n* outcomes is conditionally dependent of two other chance nodes, so the underlying table for this node will be k*m*n, where *n* is number of outcomes from a particular node, *m* and *k* are number of outcomes for these two nodes, linked to the first one. All chance nodes in the influence diagram form a belief network. In the belief network probabilistic inference is estimated by the Bayesian equation (see Equation 2 in Chapter III). According to Nielsen and Jensen (1999), influence diagram is a belief network augmented with decision and utility node(s) (NIELSEN and JENSEN, 1999).

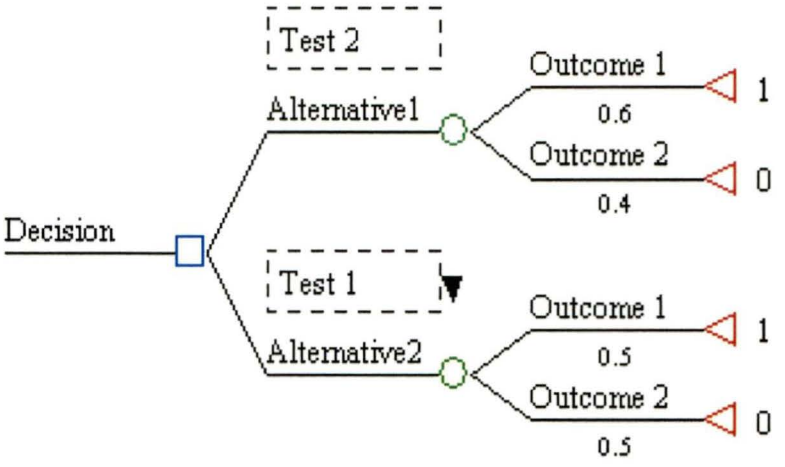
Belief (causal probabilistic or Bayesian) networks allow qualitative knowledge (structure of a problem) and quantitative knowledge, derived from case databases, expert opinion and literature to be exploited in the construction of decision support systems for diagnosis, treatment and prognosis (ANDREASSEN et al., 1999).

Decision tree

The decision tree is a graphical description of a sequential decision process. Evaluation begins at the terminal nodes and progresses backwards to the decision node. At each chance node, a value is a summary of the weighted average of the values of its possible outcomes. The strategy with the highest (or lowest) expected value is the strategy of choice. Each branch (or arc) leads to a terminal node. At terminal nodes the process stops, and the utility associated with a terminal state can be evaluated. Branches could lead to a chance node (usually represented as a circle), where the result of the event (e.g. test in Figure 1) is

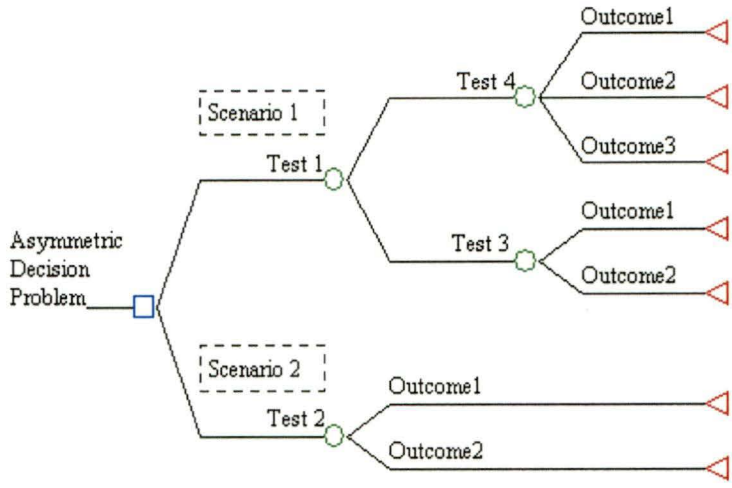
uncertain. The arcs could also terminate at nodes that either represent additional decision or terminal events.

Figure 1. Example of the generic decision tree structure (symmetric)



The order in which the nodes are traversed from left to right is the sequential order in which decisions are made and/or outcomes of chance events are revealed to the decision maker. Decision trees are easy to understand and easy to solve. If a variable is not relevant in a scenario, a model structure simply does not include it. Decision trees are symmetric if all alternatives have the same structure, or asymmetric, as shown in Figures 1 and 2. Use of decision trees is however usually limited to small problems due to the exponential growth of the representation. Conditional independence is not explicitly represented in decision trees (BIELZA and SHENOY, 1996).

Figure 2. Example of the generic decision tree structure (asymmetric)



Influence diagram

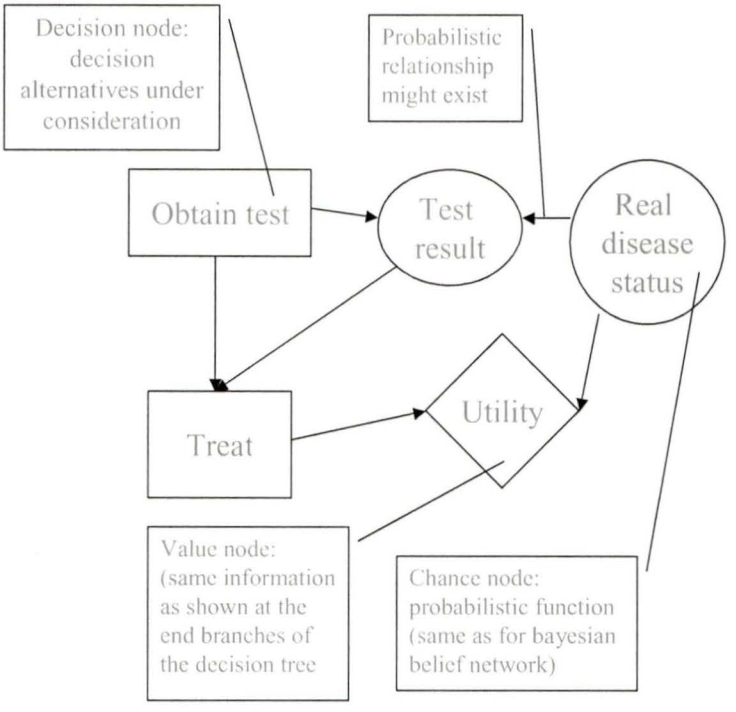
Influence diagrams were introduced as a formalism to model decision problems with uncertainty (DITTMER and JENSEN, 1997). An influence diagram can be converted into a decision tree. This approach is also used for solving influence diagrams (QI and POOLE, 1995).

Influence diagrams serve as a powerful modeling tool for symmetric decision problems. When formulating a decision scenario as an influence diagram, a sequential ordering of the decisions variables is required. No barren (unconnected) nodes are specified by the influence diagram since they have no impact on the decisions (NIELSEN and JENSEN, 1999).

To illustrate influence diagrams an example is shown on Figure 3. It contains two decision nodes (rectangles) that represent a choice to obtain a diagnostic test or to prescribe a treatment. Test results (a chance node) depends on the real disease status. Depending on test results, different treatment options may be prescribed (a decision node "Treat").

The set of value nodes (drawn as diamonds) defines a set of utility functions, indicating the local utility for a given configuration of the variables in their domain. The total utility is the sum of the local utilities. In the current example only one utility node is present. Utility is calculated on the basis of real disease status and treatment prescribed. The evaluation is performed according to the maximum (minimum) expected utility principle.

Figure 3. Example of influence diagram elements



An influence diagram representation of a problem is specified at three levels (graphical, functional, numerical). At the graphical level, a directed acyclic graph[4] displays decision variables, chance variables and information constraints. At the functional level, the structure of the conditional distribution is specified for each node. At the numerical level, the numerical details for the probability distributions and the utilities are specified. The size of an influence diagram graphical representation grows linearly with the number of variables.

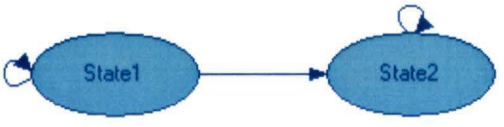
Influence diagrams are intuitive to understand and encode conditional independence relations (BIELZA and SHENOY, 1996).

Influence diagrams are less user friendly than decision trees to represent asymmetric decision problems (NIELSEN and JENSEN, 1999).

Markov models

Markov models are useful when a decision problem involves repeated events and the timing of events is important. The model assumes that the patient is always in one of a finite number of states of health referred to as Markov states. All events of interest are modeled as transitions from one state to another. Each state is assigned a utility, and the contribution of this utility to the overall prognosis depends on transition from or the length of time spent in the state. The time horizon of the analysis is divided into equal increments of time (so-called Markov cycles). During each cycle, a patient may make a transition from one state to another or to itself. On the state-transition diagram each state is represented by a circle (see Figure 4). Arrows connecting two different states indicate allowed transitions. Arrows leading from a state to itself indicate that the patient may remain in that state in consecutive cycles. A state transition diagram can be easily transformed into a decision tree representation form (see Figure 5).

Figure 4. State-transition diagram of Markov model



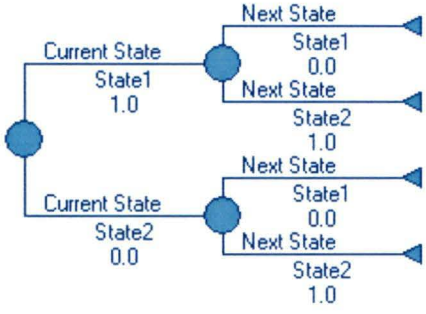
The length of the cycle is chosen to represent a clinically meaningful time interval. The utility that is associated with spending one cycle in a particular state is referred to as the

---

[4] A directed acyclic graph is a directed graph where no path starts and ends at the same node.

incremental utility. Utility accrued for the entire Markov process is the total number of cycles spent in each state, each multiplied by the incremental utility for that state. The probability of making a transition from one state to another during a single cycle is called a transition probability. The Markov process is defined by the probability distribution among the starting states and the probabilities for the individual allowed transitions. A transition matrix can be represented by n*n transition probabilities where n is number of states. Absorbing states are states that the patient cannot leave (e.g. death, etc) (SONNENBERG and BECK, 1993). A transition probability can be expressed as a single value(direct way) or as a net probability (indirect way). If it is difficult to specify the transition directly, the net probability can be modeled using chance nodes and the probabilistic influences among them (LEONG, 1998).

Figure 5. Decision tree representation of Markov model



Two different types of Markov model can be characterized by the form of the transition probabilities. A special type of Markov process in which the transition probabilities are constant over time is called a Markov chain. This has distinct analytical advantages since the probability of being in a particular state at a particular point in time can be calculated simply by raising the transition matrix to the power of the appropriate cycle. The more general Markov models, where transition probabilities can vary over time, are known as a time-dependent Markov process (BRIGGS and SCULPHER, 1998).

An important limitation of the Markov model is that the probability of moving out of a state is not dependent on the states a patient may have experienced before entering that state (so-called Markovian assumption) (BRIGGS and SCULPHER, 1998).

The accuracy of Markov model results depends on the accuracy of the estimates for the transition probabilities between different states of the model. They could be derived from cohort studies, which however could be subject to selection bias. The precision of an estimate is directly related to the number of person-years of observations for the cohort. The model uses inputs such as the probabilities of eventually dying with different stages of disease (absorption probabilities) or the mortality rates from other causes (BLACK et al., 1997).

Dynamic influence diagrams

Leong (1998) published a work on representation and solving clinical problems as dynamic influence diagrams. In a dynamic influence diagram, similar elements to other modeling approaches can be found. The dynamic influence diagram has decision nodes, chance nodes and the value nodes (like influence diagrams) and also state variable nodes. Arcs between chance nodes represent conditional dependence (LEONG, 1998).

This researcher has made an analysis of dynamic decision modeling approaches and developed an integrated decision language (DynaMoL) with four components (dynamic decision grammar, graphical representation convention, the mathematical representation and a set of general translational techniques). She has also implemented this language in a prototype and described several medical domain problems. The prototype has been further developed in a software application for decision analysis modeling ("ReasonEdge Modeler"). This software has 3 views of a model structure. The state transition diagram represents all states of the model, which can influence expected utility. Links between different states

represent allowed transitions of patients between states. Utility is referenced to "being in the state" or as "transition from one state to another". Transitional probability can be represented as a belief network or decision tree (LEONG, 1998, LEONG and CAO, 1998).

The author has shown the DynaMoL framework as a platform for automatic derivation for numerical parameters, supporting knowledge based model construction and automated knowledge acquisition from multiple knowledge sources (CAO et al., 1998, CAO and LEONG, 1997, LAU and LEONG, 1999, WANG and LEONG, 1998).

Model evaluation

Structural level

Evaluation of a model starts from the evaluation of how the model structure represents the clinical problem. The clinical problem can be described as a text or represented graphically. In decision modeling each node represents some kind of event or diagnostic test. The decision model should contain a necessary set of elements to represent the clinical problem and to answer research questions.

Functional Level

Links between chance nodes represent the nature of the relationship between the events. The conditional dependency can be visually evaluated with a belief network. The same task becomes difficult when only a decision tree representation of the decision model is used. The conditional dependencies between chance nodes are hidden behind the numbers on the branches that emanate from these chance nodes. Decision model alternatives represent scenarios relevant to clinical practice.

Numeric level

Briggs (2000) recognized three categories of uncertainty, which could be generally applied to the numeric level:

1) uncertainty relating to observed data inputs. Typically, confidence intervals might be presented, the size of which depends not only on sample size, but also on within sample variability;

2) uncertainty relating to extrapolation. This includes data generalized from other settings, as well as data modeled using epidemiological models or regression;

3) uncertainty relating to data analytic methods.

To deal with these sources of uncertainties, careful selection of numeric information should be done in order to specify the probabilistic relationship between events (BRIGGS, 2000).

The accuracy of a decision model depends on the accuracy of the estimates for the transitional and conditional probabilities used in the model (BLACK et al., 1997). Transition between Markov states is defined by the logical temporal relationship between conditions of the patient represented as Markov states (e.g. no transition from state "Dead" to state "Alive").

If a transitional probability is not available, the parameter might be estimated through conditional probabilities using a belief network. For example, transitional probability between state 1 (e.g. "healthy patient") and state 2 (e.g. "disease") is unknown. But the model developer could estimate the probability of developing the "disease" if patients were exposed to some factor, and the prevalence of this factor in the population is known. Transitional probability is calculated using the Bayesian equation (see Equation 2 in Chapter III). A belief network for transitional probabilities can be useful also for conducting

sensitivity analysis on some parameters within the belief network (LEONG, 1998, NIELSEN and JENSEN, 1999).

Small samples or short follow-up gives confidence intervals that are large relative to the transition probabilities. Patients selected retrospectively are often chosen on the basis of availability (BLACK et al., 1997).

Sensitivity analysis

A sensitivity analysis involves systematically examining the influence of uncertainties in the variables and assumptions employed in an evaluation on the estimated results. This method is included to ensure the significance of the obtained results. Krahn et al.(1997) have named a sensitivity analysis as "the decision analyst's version of statistical hypothesis testing" (KRAHN et al., 1997).

One way sensitivity analysis systematically examines the impact of each variable in the study by varying it across a plausible range of values while holding all other variables in the analysis constant at their "best estimate" or baseline value. Even if an analysis is robust to changes within a single variable, it may require multi-way sensitivity analysis (KRAHN et al., 1997).

An analysis may show insensitivity to the one-way or multi-way changes. In this case, the model is robust. At the other extreme, the analysis may be sensitive to small changes in variables. Thus a critical judgment, based on the sensitivity analyses and an evidence quality, should help the decision about an optimal strategy or an alternative. A sensitivity analysis also determines which variables require further empirical evaluation (KRAHN et al., 1997).

The term "sensitivity analysis" in decision modeling is distinct from "sensitivity" for diagnostic performance evaluation. During "sensitivity analysis" the decision maker

evaluates the robustness of the results while "sensitivity" of a diagnostic performance measure estimates the ability of a diagnostic test to detect positive cases. These two measures were both used in the current work.

Utility calculation techniques

Utility in decision analysis modeling is a quantitative evaluation of outcome from a decision makers perspective. Expected utility of a particular alternative is a weighted average utility of all possible outcomes of a probabilistic situation. There are two methods used for expected utility calculation. The first one, referred to as a roll back procedure, is based on the multiplication of the utility associated with a particular outcome and the probability of this outcome to happen. Another method uses a single hypothetical patient run through the model with Monte-Carlo simulation, as follows. The patient starts in one of the initial health states in the model. Each subsequent transition in the Markov model is simulated by random draws from a uniform distribution. The transition is made if the value of the draw exceeds the point probability estimate for the transition. A patient entering the process is "followed" through the simulation until an "absorbing" state or stopping criteria is reached. Running this simulation many times results in a probability distribution of the relevant outcome for the individual (CHER and LENERT, 1997). Monte-Carlo simulation allows to provide confidence intervals around estimates and may produce a more realistic estimate of uncertainty (BRIGGS, 2000, GROVER et al., 2000a).

Information sources

Three sources of evidence are considered for decision analysis (literature, human experts and original raw data from clinical practice). The first source is usually called external evidence, the second and the third one are referred to as internal sources of evidence

(BUXTON et al., 1997, MEHTA et al., 1998). This information can be used to define the structure of the decision model and also for probabilities and outcome assessment. Literature, as an evidence source, means any summarized statistical data from relevant studies (LEHMANN et al., 2000). Human experts and domain knowledge about a specific subject (e.g. "prostate cancer") might be used for decision modeling as the source of structural assumptions.

Evidence based medicine (EBM) principles specify the review and assessment of all available evidence and synthesis of information with assessment of results validity. Cochrane Collaboration Centers[5] have developed a systematic review method of scientific publications. By this method, all possible evidence is to be evaluated by 4 major criteria (patient population, intervention, outcomes and control/comparison) in order to select publications relevant for a current study. Evidence from relevant studies using aggregation techniques forms our study data.

A decision tree as well as a belief network represents a patient population, that undergoes specific interventions and has specific outcomes. Three attributes (Population, Intervention and Outcome) were applied to select evidence for every chance node. Table 2 shows definitions and comparison of meaning for each of these attributes.

[5] http://www.cochrane.org/ Accessed November 28, 2003

Table 2. Definitions and meaning of each attributes of chance nodes.

Attribute	Cochrane Collaboration Centers	Decision modeling
Population	Inclusion and Exclusion criteria	Patient group defined by the parent chance nodes
Intervention	Intervention used in study to test the difference between groups	Chance node itself as a representation of intervention (diagnostic test or treatment)
Control/ Comparison	Shows how many similar groups of patients were studied (two for controlled studies ; one for non-controlled studies)	Beneficial approach included for evaluation
Outcome	Criteria to study difference between groups (e.g. mortality, specific events, so called end-points)	Outcomes which represent groups of patients by some events

Some limitations were observed. It is difficult to detect the difference in screened and non-screened population mortality due to limited data sets (limited follow-up period). Clinical practice usually provides datasets on restricted populations due to variations of clinical protocols and cost minimization in health care. In general there are few biopsy data when tPSA is less than 4 μg/l. Fleshner et al. (2000) have shown significant difference in using biochemical markers between Canada and USA for identification of new cancers and radical treatment indications (FLESHNER et al., 2000).

When dealing with heterogeneous data from the literature, Lehman et al. (2000) has listed 3 approaches of evidence synthesis: 1) listing evidence for individual probabilities; 2) summarizing evidence across probabilities; 3) integrating the pooled evidence for individual probabilities into the decision model. The probabilities can be summarized in 3 ways: by averaging, by averaging weighted by sample size (pooled), and by meta-analysis (LEHMANN et al., 2000).

Variability of sources and heterogeneity of data can make it difficult to draw the qualitative and quantitative conclusions. Methodologies developed for assessing evidence in evidence based decision making and those applied to decision analysis are essentially the same.

Results representation

Decision analysis studies are not so widespread as usual studies. Redelmeier et al. (1997) have therefore highlighted the need to pay attention to decision analysis representation because reviewers and readers might have difficulties with understanding all aspects of such study design (REDELMEIER et al., 1997).

Defining the levels of model representation (structural, functional and numeric) was suggested by Bielza and Shenoy (1996), and Matzkevich and Abramson (1995). The meaning of these levels is discussed previously in this chapter (BIELZA and SHENOY, 1996, MATZKEVICH and ABRAMSON, 1995).

Criteria for reporting clinical studies can be applied to decision analysis study results. Briggs and Sculpher (1998) have suggested to use more descriptive statistics when reporting expected utilities accompanied with interval estimates.

Authors have reported skewness of cost and cost-effectiveness data and have suggested to include all variables to the sensitivity analyses of a model (BRIGGS and SCULPHER, 1998, Briggs, 2000).

Evaluation of Utility

Several types of economic analysis (cost-minimization, cost-efficacy, cost-effectiveness, benefit-cost, etc) might have been distinguished in the literature. Scenarios are evaluated according to ratio of two parameters (e.g. cost/effectiveness). Ratio serves as a single utility. During cost-minimization analysis, scenarios are compared by costs only. It might be applied when technologies are equally effective.

According to Kessler (1997), and Nielsen and Jensen (1999), "effectiveness" refers to those outcomes and response rates achieved in clinical practice and depends on a number of factors, including patient variation, resources and structures, physician variation, severity

of disease, concomitant therapy, and patient compliance (KESSLER, 1997, NIELSEN and JENSEN, 1999). Gould and Birkmeyer (1999) describe "efficacy" as a performance measure of an intervention for a given health problem under the "ideal conditions" of an investigation whereas "effectiveness" is based on results obtained under usual conditions of clinical care for a particular group (GOULD and BIRKMEYER, 1999). In comparison to efficacy and effectiveness, benefit is expressed in monetary terms as an economic achievement for a given population (e.g. cost reduction).

A cost-effectiveness analysis calculates the cost of producing a desired effect (often measured in units of quality adjusted life-years) but offers no judgment regarding relative worth or willingness to pay. It could be useful when effectiveness of the compared technologies are different, activities with the same aim and measure of effectiveness are compared.

Cuzin et al. (1998) suggests examples of efficacy used in decision modeling of cancer problems (total number of cancer cases detected or per stage; expected life gain). The following ratios are possible (cost per screened person, cost per detected cancer, cost per treatable cancer; cost per local stage of cancer; cost per saved life) (CUZIN et al., 1998).

A benefit-cost analysis provides a view on the assessment of the worthiness of funding a project. A cost-effectiveness analysis may compare different options, once a decision to proceed has been made (KRISTENSEN et al., 2001, LITTRUP et al., 1994a, NIELSEN and JENSEN, 1999).

Influencing factors

Cost per cancer detection with time (marginal cost) increases exponentially when serial screening tests detect progressively fewer tumors. Digital rectal examination had the lowest marginal cost in the first 2 years. The low sensitivity of digital rectal examination for

subsequent cancer caused its marginal cost to increase rapidly. Thus a tPSA oriented approach becomes less costly for screening programs lasting more than 3 years (BENOIT and NASLUND, 1997).

The following end-points could also be used in the prostate cancer domain: overall survival, disease specific survival and cancer recurrence. These end-points require different follow-up study periods. The longest one is for overall survival, and the shortest is for cancer recurrence (bone scan detected or biochemical recurrence detected by tPSA level changes). A cancer recurrence is more attractive for evaluation of cancer treatment rather than a mortality related end-point. The disease specific mortality could be used to eliminate the effect of mortality from other reasons. Using such a parameter would be important for a group of aged patients, where the partial role of other mortality reasons is higher (GARNICK and FAIR, 1996a, GARNICK and FAIR, 1996b).

Available prostate cancer models

This section provides comparative analysis of 26 prostate cancer decision models in the last decade (CANTOR et al., 1995, COLEY, 1997a, COLEY, 1997b, DRAISMA et al., 2003, ELLISON et al., 2002, ETZIONI et al., 1996, FLEMING et al., 1993, GOTTLIEB et al., 1996, GROSSFELD, 2000, GROVER et al., 2000a, GROVER et al., 2000b, GUSTAFSSON et al., 1995a, HILLNER et al., 1995, JAGER et al., 2000, KATTAN et al., 1997, KRAHN et al., 1994, LAUNOIS, 1992, LITTRUP et al., 1994a, LITTRUP et al., 1994b, MENG and CARROLL, 2000, MOLD et al., 1992, OGAWA and KATO, 1998, ROSS et al., 2000, SEIDENFELD et al., 1999, WOLF et al., 1993b, WOLF et al., 1995, YOSHIMURA et al., 1998).

This analysis helped to elaborate the needs for a new model and define an optimal modeling approach for prostate cancer. The analysis highlights technical aspects of prostate

cancer decision models in respect to the problem studied. The publication date of the studies for all models starts from 1990, which corresponds to the period of introducing tPSA in clinical practice. A detailed comparison table of published models is available in the appendix. The term non-Markov model in the following section corresponds to the decision trees and influence diagrams.

Decision modeling for prostate cancer domain

Chodak (1993) highlighted the importance of decision analysis because it offers a method for understanding the implications of alternative screening strategies and provides a basis for deriving the most reasonable approach to prostate cancer screening (CHODAK, 1993).

In 1998 an expert group from France published a report on the evaluation of evidence relevant to prostate cancer screening including an evaluation of published decision models for prostate cancer. As part of the assessment they proposed the following axes for decision analysis model evaluation (adopted perspective; evaluation type; valued strategy; population; method of cost evaluation (direct or indirect costs), number of strategies, integrated parts (e.g. treatment, staging), data source (pro/retrospective) (CUZIN et al., 1998). Another framework for the evaluation of cost evaluation studies was published by the British Medical Journal (BMJ) in 1996 (DRUMMOND and JEFFERSON, 1996).

In this current work a synthesis of these two frameworks has been used to review previously published models of prostate cancer.

Table 3. Evaluation frameworks for decision modeling studies

Summary of BMJ economic evaluation guidelines (from Drummond and Jefferson, 1996)	Summary of Report on prostate cancer models by Cuzin et al., 1998	The adapted framework for current work
❑ Study design - Study question - Selection of alternatives - Form of evaluation ❑ Data collection - Effectiveness data - Benefit measurement and valuation - Cost data - Modeling ❑ Analysis and interpretation of results - Adjustment for timing and costs of benefits - Allowance for uncertainty - Presentation of results	❑ Adopted perspective ❑ Evaluation type ❑ Valued strategy ❑ Population ❑ Method of cost evaluation (direct or indirect costs) ❑ Number of strategies ❑ Integrated parts (e.g. treatment, staging) ❑ Data source (pro/retrospective)	❑ Type of evaluation (retro-prospective, cost-benefit, cost-efficacy, cost-effectiveness, etc) ❑ Objectives ❑ Population ❑ Methods (decision tree, influence diagram, Markov model, Monte-Carlo simulation, etc) ❑ Alternatives ❑ Phases (screening, staging, treatment) ❑ Cost estimation method ❑ Results

A detailed review of published models of prostate cancer using the adapted evaluation framework is given in the appendix.

Research questions studied

The research questions presented in the published models can be arranged into the following groups:

i. non-screening vs. screening (CANTOR et al., 1995, COLEY, 1997a, COLEY, 1997b, DRAISMA et al., 2003, GOTTLIEB et al., 1996, HILLNER et al., 1995, LITTRUP et al., 1994b);

ii. some screening strategies are better than others (if screening is beneficial itself) (ELLISON et al., 2002, ETZIONI et al., 1996, GUSTAFSSON et al., 1995a, KRAHN et al., 1994, LAUNOIS, 1992, ROSS et al., 2000);

iii. results are sensitive to changes in diagnostic sensitivity/specificity of tests (LAUNOIS, 1992);

iv. use of some stage determination tests (e.g. pelvic lymph node dissection, computer tomography (CT) scan, medical resonance imaging (MRI) scan are only advisable for certain groups of patients (JAGER et al., 2000, MENG and CARROLL, 2000, WOLF et al., 1993b, WOLF et al., 1995);

v. choice of treatment options (CANTOR et al., 1995, FLEMING et al., 1993, GROSSFELD, 2000, HILLNER et al., 1995, KATTAN et al., 1997, OGAWA and Kato, 1998, SEIDENFELD et al., 1999, YOSHIMURA et al., 1998);

vi. results are sensitive to effectiveness criteria of treatment(CANTOR et al., 1995, COLEY, 1997a, COLEY, 1997b, KRAHN et al., 1994);

vii. survival forecast (GROVER et al., 2000a, GROVER et al., 2000b).

Scenarios

In decision making, a scenario, or alternative, refers to a set of conditions (population, other tests or interventions) under which a medical technology is tested. In decision modeling of healthcare problems a scenario is a part of the decision model, which represents a set or sequence of diagnostic tests, interventions and outcomes. Scenario variation and number of scenarios in prostate cancer models correlates with the research questions studied. Authors often use more than two scenarios to evaluate.

Most alternatives created are for screening and stage determination. These alternatives include diagnostic test combinations. Due to limited data availability for longitudinal models the number of scenarios for non-Markov models is higher than for Markov models (GOTTLIEB et al., 1996, LAUNOIS, 1992, LITTRUP et al., 1994b, WOLF et al., 1993a, YOSHIMURA et al., 1998). Grover et al. (2000) used a Markov model with

only one scenario, to forecast the 10-years disease specific survival for prostate cancer (GROVER et al., 2000a).

Authors usually did not adjust the diagnostic performance of screening scenario to the same level within one model. This may cause biased results while comparing expected utilities for such alternatives.

Data source

Most of the time authors have used retrospective data of clinical practice or other research studies. Only few have used a prospective study design (GUSTAFSSON et al., 1995b, LITTRUP et al., 1994b). Only one author published prospective study results on economic implications after 5 annual prostate cancer screenings (METTLIN et al., 1991).

Published scientific studies were the most used probabilistic data source. Many authors created Markov models solely with the data from published studies (CANTOR et al., 1995, COLEY, 1997b, JAGER, 2000, KRAHN et al., 1994, MENG and CARROLL, 2000, MOLD et al., 1992, WOLF et al., 1995). In comparison to them Hillner et al. (1995) used data from a clinical study database, but published studies were used to estimate range for sensitivity analysis on probabilistic parameters for the model (HILLNER et al., 1995). The most recent models are developed with combination of data from scientific studies and information from national level cancer registers (GROVER et al., 2000a).

Modeling approach

Using Markov or non-Markov model highly depends on the nature of the problem studied (e.g. screening, stage determination or treatment). Treatment scenario can be compared using a decision tree (e.g. life expectancy is known) or Markov model (life expectancy is modeled using survival rate). Evaluative studies for staging diagnostic tests

(e.g. tPSA, Gleason Score, imaging tests, etc) are often modeled as a decision tree. Screening scenarios were modeled as combinations of diagnostic tests (e.g. DRE, tPSA, etc). Only one cost-effectiveness study using complexed PSA for screening has been published (ELLISON et al., 2002).

Authors often have included treatment outcome of the screening models in order to assess cost and quality-of-life for population undergoing serial prostate cancer screening. However using quality-of-life parameters for assessing needs for a screening at all was criticized by Krahn et al. (1994). Some authors have stated that the result "to screen" or "not to screen" is highly sensitive to utility based on quality-of-life parameters (KRAHN et al., 1994). Thus evaluation of a screening program depends on the post-screening outcomes. Quality of life differences multiplied by many years can significantly affect the end result.

In one study a formula was used instead of graphical modeling approach. Authors created a complex formula to calculate expected utility for a serial screening program. (LITTRUP et al., 1994a).

Utility estimation

Decision modeling has been used for almost all studies where a cost estimation of prostate cancer screening has been done. One approach for cost estimation is about saving on screening strategy modifications. An example is the Woodrum (1998) study about using percent free PSA for a tailored biopsy approach (WOODRUM, 1998). A small number of prospective studies on consecutively screened patients over several years are available. Models, which adjust costs in accordance with the shifts in such parameters as cancer detection rates and tumor stage that occur with serial screening, are needed. Another approach is cost of finding one case of prostate cancer. Candas et al. (2000) estimated the difference of cost per one cancer diagnosed for the first and follow-up visits as $2'420 CAD

and $7'105 CAD respectively (CANDAS et al., 2000). The difference in cost of diagnosed prostate cancer at the first and follow-up visits due to decreasing prevalence of the disease suggests modification of diagnostic strategy. The increase of the cut-off also increases the specificity.

Cost/effectiveness parameters have been used most of the time. However costs were estimated in different ways. Some researchers took the costs from national databases (COLEY, 1997b, LAUNOIS, 1992), the others estimated costs based on a third party perspective (e.g. Medicare, etc) (GROVER et al., 2000b, HILLNER et al., 1995, JAGER, 2000).

Evaluation of screening program has been done using life-quality parameters. Most of the modeling studies for prostate cancer screening evaluation have used an effectiveness parameter which was defined as a preference of outcome rated by patients' or by doctors' feelings. Grossfeld (2000), Mold et al. (1992), Meng and Carroll (2000) selected a panel of experts. On the other hand Kattan et al. (1997) and Wolf et al. (1993) selected patient groups, while Cantor (1995) determined utility with 10 married couples (CANTOR et al., 1995, GROSSFELD, 2000, KATTAN et al., 1997, MENG and CARROLL, 2000, MOLD et al., 1992, WOLF et al., 1993a). In the Krahn et al.(1994) study the effect of complications on quality of life were based on estimates from a small group of 10 urologists, radiologists, and oncologists (KRAHN et al., 1994). The quality of life adjustments assigned to complications of treatment have not been based on standardized, validated questionnaires.

With this new model we will estimate strategies for an annual prostate cancer screening using a new biochemical marker (cPSA) and a choice of strategies with respect to the cost of a 5 years screening program. A model of the natural history of prostate cancer that

includes prostate cancer screening and diagnosis as well as patient survival may generate new evidence in the prostate cancer domain.

II. Research hypotheses and objectives

Research hypothesis

The review of prostate cancer studies results suggest that the implementation of prostate specific antigen (PSA) based screening might provide survival advantage for patients with prostate cancer. Screening for prostate cancer should lead to early detection, where the latter can identify smaller cancer and enable treatment earlier. Treatment of early cancer should be associated with improved survival. Hence this might lead to the demographic changes: screening -> early detection -> smaller cancers -> treat earlier -> improvements in survival -> demographic changes. The study of this should benefit from a modeling approach.

1. Decision modeling might be used as an instrument for integration of evidence in the evaluation of prostate cancer.

2. Quantitative results of an integrative decision model may serve as new evidence in the problem domain by answering different questions depending on model structure and data availability.

Research objectives

1. To investigate the use of decision modeling to enable integration of evidence in the evaluation of prostate cancer.

2. To develop a decision model for screening/stage determination/treatment of prostate cancer and specifically evaluate

- screening strategies using a new biochemical screening marker
- integrated model behaviour with change in model parameters, particularly with respect to cost and survival.

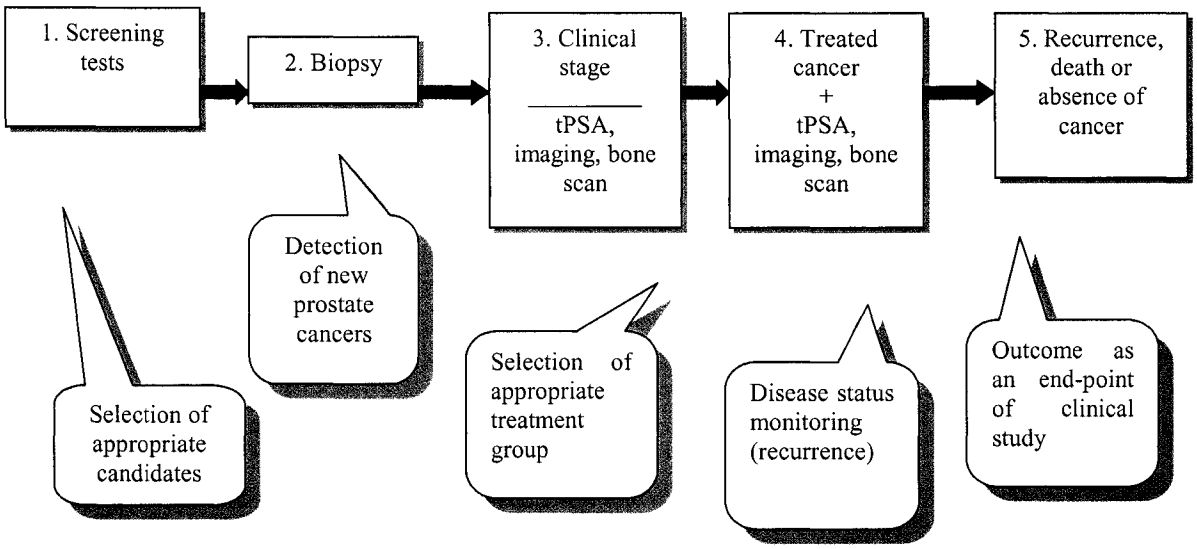
III. Materials and methods

The first subject to be discussed is the conceptual representation of the prostate cancer domain that will be represented as a decision model. The main phases are identified and models are developed initially separately for each phase. The description of each phase is based on the principles for decision analysis model development described by Leong (1998). We previously described the prostate cancer domain in Chapter I: "Prostate cancer". The data sources are identified for each of the phases. Literature analysis is based on the principles of a systematic review. Clinical data are described. Ethical considerations are given on clinical data use. The modeling approach and software are proposed and the criteria for model evaluation determined in relation to the study objectives. The calculation approaches for transitional and conditional probabilities and expected utility are given based on the study objectives and software functionality.

Conceptual representation of the clinical problem (structural level of model)

The natural history of prostate cancer is represented schematically to reproduce all major events that could happen during the life of a man consequent to prostate cancer screening and diagnosis.

Figure 6. Schematic representation of the natural history of prostate cancer



The model was created in three phases (screening and diagnosis, staging, treatment outcome). The first phase model represents prostate cancer screening and diagnosis of new prostate cancer cases. Being involved in the screening and diagnosis if necessary is one repeated state, having newly diagnosed prostate cancer case is another state within the first phase model. Serial prostate cancer screening during a patient's life can be represented as a chain of events repeated at intervals. During screening the patient can come back into the population of presumed healthy people (transition 1 → 1) or could become a candidate for biopsy . This biopsy can reveal a newly diagnosed prostate cancer (transition 1 → 2) or the patient might be eligible for next screening if no abnormality was detected (transition 1 → 2 → 1).

Staging is a diagnostic process of how far the cancer is spread in order to provide an appropriate recommendation for treatment. Clinical stage relates to the findings on DRE, which provides information on primary tumor volume and whether the disease has spread locally out of the prostate. Imaging is able to see whether cancer has spread outside the prostate gland and help determine the clinical stage of the cancer. The pathologic stage is

determined after radical prostatectomy (surgical removal of the prostate gland and surrounding structures). A pathologic stage cannot be definitively determined for men who do not undergo this surgery (transition 3 → 4).

Treatment is provided to prostate cancer patients according to which treatment group they are assigned (transition 3 → 4).The major events in the post-treatment period are the recurrence of cancer and death. Death can be due to prostate cancer or due to other reasons. Recurrence of prostate cancer can be diagnosed by clinical signs, imaging tests or with monitoring of tPSA after treatment (transition 4 → 5).

Information (structural assumptions and transitional probabilities) from these three models are used to create a single model from screening to treatment outcome. The treatment outcome phase (the third phase) should enable evaluation of the choice of diagnostic strategy on treatment outcomes. In our study, clinical experts (urologist, radiologist) were interviewed for model structure validation.

Information sources (numerical level of model)

The data for model construction were derived from a systematic review of published studies and also based on raw data obtained from a multicentre study of new biochemical screening markers of prostate cancer. The latter source was relevant for the screening, diagnostic and staging components but not the treatment component.

Probabilistic information from the Bayer clinical study data

The main source of probabilistic information for the screening and diagnosis stage was data obtained from the "Multicenter clinical performance of the Bayer Immuno1 Complexed PSA assay in the screening population" (1202 patients), which is referred to later as the "Bayer clinical study". This study was conducted in 2001 and the data were used for

cost-effectiveness analysis on new biochemical markers, published in 2002 (ELLISON et al., 2002). The purpose of this study was the evaluation of diagnostic performance of complexed PSA for prostate cancer diagnostics. The study includes retrospective (no greater than 50 subjects from each of seven sites) and prospective (minimum 150 patients for each of seven sites in the United States and Austria) follow-up of subjects scheduled to undergo needle prostate biopsy. The most important inclusion criteria were the following: no personal history of prostate cancer or trans-urethral resection; known age, patient history, race; tests of DRE; tPSA, cPSA, fPSA, transition zone and total prostate volume measurement have to be done. Exclusion criteria were certain medication use, food supplements, blood sample taken after biopsy, bad storage condition of specimens. The permission to use the data was obtained – see below in the ethics section. The demographic information about the study population is shown in Table 20 in Chapter IV Results.

Probabilistic information from the literature

The methodology for systematic review consists of several consecutive steps. The first step is to identify research questions that need to be solved using systematic review and/or meta-analysis (e.g. several studies with similar research question were published, and there are needs for a more precise answer on that research question). This is followed by evidence selection, evidence evaluation, and summarizing the evidence.

Systematic review principles were used to select appropriate publications in online bibliographic databases (CancerLit, Current Content, Medline, PreMedline). CancerLit is a bibliographic database produced by the US National Cancer Institute. The major focus for all records is a cancer therapy. Some information in CancerLit is cross referenced in the Medline database. Approximately 200 core journals contribute a large percentage of the records. The Current Contents Search database provides access to the tables of contents and

bibliographic data from current issues. Medline is the U.S. National Library of Medicine's[6] bibliographic database that covers the fields of medicine, the health care system, and the preclinical sciences. Medline contains bibliographic citations and author abstracts from more than 4,600 biomedical journals published in the United States and 70 other countries. PreMedline the National Library of Medicine's (NLM) in-process database for Medline, provides basic information and abstracts before a record is indexed with MeSH heading and added to Medline (from the OVID official web site[7]).

To be included in this analysis, studies should be relevant to the decision model in general and a decision model element in particular. Articles were excluded if the data were inadequate to satisfy model needs. For example, inclusion criteria for a decision model element representing the tPSA screening test were the following: measurements were done using tPSA assay, and tPSA values were available within the article full text. Exclusion criteria were based on diagnostic test outcomes. For example, if outcomes for the decision model element representing tPSA screening test were "more than 4 μg/L" and "less than 4 μg/L" and tPSA values in target values were presented as "more than 5 μg/L" and "less than 5 μg/L", this article was rejected (Toubert et al., 1990).

The search strategy was based on a particular element of the model. E.g. for the first phase model (screening and diagnosis) keywords included "prostate cancer" and "screening". Published studies on prostate cancer screening were used to find appropriate alternatives. Articles for elicitation of probabilities for each node in the belief network have been selected by the conditional relationship between the nodes. For example, a "biopsy" node is connected with "tPSA" and "DRE" nodes. Hence keywords for a search strategy included "biopsy", "tPSA" and "DRE" (see Table 4 for more examples). A target article

---

[6] http://www.nlm.nih.gov Accessed November 28, 2003

should provide probabilistic information to specify the conditional relationship between these three nodes, otherwise it was excluded as inadequate to a decision model.

Table 4 Keywords for a search in bibliographic databases

Any intervention or diagnostic test, represented by a decision model element	Keywords for a search in bibliographic databases
Digital rectal examination	((DRE) OR (digital rectal examination)) AND (prostate cancer screening)
Prostate specific antigen	((PSA) OR (prostate specific antigen)) AND (prostate cancer screening)
Prostate biopsy	(prostate cancer) AND (biopsy)
Bone scan	(prostate cancer) AND (bone scan)
Pelvic lymph nodectomy	(prostate cancer) AND ((pelvic lymph nodectomy) OR (PLND))
Prostate cancer treatment groups survival	(prostate cancer) AND (treatment) AND (survival)

The sum of probabilities weighted by study population size was used (LEHMANN et al., 2000). By using this approach, outcome probabilities (p_i) and study size (W_i) were multiplied for each study. An average of these values results in the probability of a particular outcome for a model (see Equation 1).

Equation 1. Formula to calculate summary probability from multiple studies

$$P = \frac{\sum p_i * W_i}{\sum W_i}$$

Review of the literature was focused on the original studies. Search was limited to English language and publication date between 1990-2002. The focus of the current study is the population after the tPSA biochemical test was measured in clinical practice. This constraint was found important because clinical characteristics of the population of men

[7] http://gateway.ovid.com. Accessed August 18, 2003

involved in screening are different before and after tPSA was introduced in the early 1990s. Heterogeneity of probabilities derived from the literature was evaluated using the Chi-square statistic.

Utility and Cost

Different perspectives (hospital, health care, third party, patient) for a cost estimation were used in published decision analysis articles on prostate cancer. Cost estimation from health care perspectives requires estimation of cost of treatment outcomes. Cost estimation from the third party's (e.g. insurance companies) or patient's (e.g. "out-of-pocket" money) perspectives is popular in United States due to prevalence of the private health care sector there. A hospital perspective was chosen. A description of different approaches used for cost estimation was provided in Chapter I.

Applying the model

Leong (1998) provided a framework for model formulation as follows: specify a problem type and evaluation criteria, define the alternative actions and the states, identify transitions directly or indirectly, and special assumptions for the actions, states and the decision parameters. A direct way of representing transition between states consists of drawing the line in state-transition diagram and assigning a probability value. As an indirect way of drawing transitions for a model, Leong (1998) uses the event variables and their relations within the underlying belief network (LEONG, 1998).

Screening and diagnosis phase (phase 1)

Patients during this phase could be in two states (general patients and newly diagnosed prostate cancer). Various diagnostic tests are provided for prostate cancer screening and diagnosis. The most popular approach is tPSA and DRE, followed by biopsy.

As alternatives, in this study, c/t PSA and f/t PSA ratio were also evaluated as possible improvement in specificity of the diagnostic program. Some authors have omitted DRE during screening. cPSA has also been suggested as an alternative for tPSA. Based on practice as described in the literature, six alternatives were proposed for study (tPSA+biopsy; cPSA + biopsy; tPSA+DRE+biopsy; cPSA+DRE+biopsy; tPSA+DRE+f/tPSA+biopsy; tPSA+DRE+c/tPSA+biopsy). Evidence for these tests ('internal evidence') was also available from the Bayer clinical study dataset.

To enable equivalent cost analysis, alternative strategies were evaluated by diagnostic accuracy parameters (sensitivity, specificity) (LAUNOIS, 1992). The cut offs for diagnostic tests were adjusted to support the same sensitivity level for all screening programs possible. Since the results of biopsy for all patients from the clinical dataset were available, ROC curve analysis was used to visually adjust diagnostic accuracy parameters. The ROC curve was plotted using MedCalc software. [8]

Stage determination phase (phase 2)

Organ-confined cancer is the most frequent diagnosis after clinical assessment of prostate cancer. In order not to miss the extraprostatic invasion or distant metastases, additional stage determination diagnostic procedures are provided. Imaging (bone-scan and/or CT scan, positron emission tomography (PET) later) is the best procedure at the present time for detection of non-organ confined cancer cases. These tests are very expensive, thus tPSA and/or Gleason Score have been suggested to pre-select patients for imaging tests. Lymph nodectomy before radical prostatectomy is used to assess the lymph node status.

---

[8] http://www.medcalc.be Accessed November 28, 2003

Clinical stage distribution (Whitmore classification) for newly diagnosed patients was available from the clinical dataset for referred patients. tPSA and GS for these patients were available from the dataset and also from the literature source. tPSA and GS are often used (separately or together) for selection of patients for imaging procedures in order to identify non-organ confined cancer. Among imaging procedures bone scan evidence is better presented in the literature in comparison to the CT scan. Literature search on PET scan articles did not provide probabilistic data which were sufficient enough to be included into our decision model.

Treatment phase (phase 3)

Treatment outcomes were studied separately for different groups of prostate cancer patients based on staging results. The first separation is done by clinical stage (confined and non-confined cancers). Confined cancer by clinical assessment could be further evaluated by imaging tests if tPSA and/or GS are over a certain cut-off. Lymph nodectomy is done before the operation for patients local stage and negative imaging results (either imaging was provided or not by stage determination protocol).

Based on treatment groups, the patient populations description was used as selection criteria for the articles in bibliographic databases. Recent randomized controlled trials (RCT) were the target for literature review. For RCTs assessing the alternative treatment approaches, preference was given to the study branch with a treatment that provided significantly better results among others.

The estimation of overall patient survival is modeled as follows. A weight of 1 is attached to each state of the model in which the patient is alive and a weight of 0 is attached to dead state. Running the model over a large number of cycles and summing the weights

across those cycles gives an estimate of the average life expectancy of the patients in terms of the model cycle length (Briggs and Sculpher, 1998).

Choice of modeling software

We have used two software application for building the decision model (ReasonEdge Modeler[9] and Data 3.5[10]). The comparison between them is shown in the table below.

Table 5. Comparison between two decision modeling software

	ReasonEdge Modeler	Data 3.5
Graphical representation of Markov model	State transition diagram, belief network and tree structure	Tree representation of Markov model
Transitional probabilities representation	Belief network	Tree structure
Compactness of representation	Very compact	Enlarged
Calculation of transitional probability from multiple sources of evidence (multi factor dependency) with using Bayes formula	With belief network	Create belief network separately and then generate tree structure
Assigning utility to the model structure	Utility can be assigned to any transition. When sensitivity analysis is performed, utility is not adjusted to the patient distribution. The only solution is to use average utility for current transition.	Utility can be assigned to any branch in tree structure, that represents state transitions. This allows adjustment of transition utility, when sensitivity analysis is performed on probabilistic parameters within decision tree
Calculation method	Rollback	Roll back and Monte-Carlo simulation. Monte-Carlo simulation, which allows to generate utilities distribution and test the significance of the difference between all alternatives

ReasonEdge was the first software choice for the development of the decision models and their optimization. At the beginning, three parts of the final model were created in ReasonEdge Modeler as two-state models. Transition probabilities were defined through belief networks. With the ReasonEdge software the user is able to assign utilities for

[9] http://www.reasonedge.com Accessed on the 3rd of Aug 2004
[10] http://www.treeage.com Accessed on the 3rd of Aug 2004

transition from one state to another. However while doing sensitivity analysis on probabilistic outcomes for chance nodes (=events) in a belief network, the utility for single transition stays unchanged. Data 3.5 software was used to integrate these three parts of the single model. With Data 3.5 software, the user can assign utility at terminal nodes as well as for every branch of the decision tree. This gives more precise results with sensitivity analysis of the model.

Transitional probabilities

Transitional probabilities were used to express the probability of different patient outcomes in the third phase Markov model. Data for this part of the model (survival of treated prostate cancer) were taken from the literature as a rate. The rate describes the number of occurrences of an event for a given number of patients per unit of time. Rate range is from zero to infinity. A probability describes the likelihood that an event will occur in a given length of time. Probabilities range from zero to 1. The probability of an event that occurs at a constant rate (r) in a specified time (t) is given by the equation $p=[1-e^{-r*t}]$. Parameters r and t could be taken from the selected study as number of events per study population and study period respectively (SONNENBERG and BECK, 1993).

The survival curve is $f=(e^{-r*t})$, where (f) is the fraction surviving at time (t) and (r) is the constant transition rate. A survival curve is a usual graphical representation for clinical study design where survival analysis was used. This graph could also be used for estimation of transitional probabilities estimation if r and t were not clearly defined in the article. At any given time the fraction that has experienced the event is equal to (1-f). Thus the curve describing the probability that the event will occur in time (t) is (1-f), or ($1-e^{-r*t}$). The probability of transition in time (t) is always less that the corresponding rate per time (t) because as the cohort members die, fewer are at risk for the transition later in the time period.

When the model is executed, for each clock cycle, the appropriate mortality rate is calculated from a formula and converted to a transition probability. The necessary rates (or probabilities) may also be stored in a table, indexed by cycle number, and retrieved as the Markov model is evaluated. Incremental utilities may vary with time.

Conditional probabilities

Conditional probabilities with belief networks were used to specify transitions in the Markov model where the indirect way of expressing transition probabilities was appropriate. This refers to the screening and diagnosis and the stage determination phases of the model (phase 1 and 2 of the model respectively).

A decision modeling software calculates probabilities for the Markov model from the belief network by Bayes' theorem.

Equation 2. Bayes theorem equation

$$P(B_i \mid A) = \frac{P(B_i) \cdot P(A \mid B_i)}{P(B_1) \cdot P(A \mid B_1) + P(B_2) \cdot P(A \mid B_2) + \ldots + P(B_k) \cdot P(A \mid B_k)} \quad \text{for } i = 1, 2, \ldots \text{or } k$$

$P(B_i|A)$ is a probability of event $[B_i]$ if event $[A]$ happens. $P(B_i)$ is a probability of event $[B_i]$ and $P(A|B_i)$ is a probability of event $[A]$ if event $[Bi]$ happens. For example, assume that a patient is provided with tPSA testing. If tPSA is over a certain cut-off, it means that a biopsy is provided. Let P(A) be a probability for tPSA over the cut-off and P(B|A) is a probability of positive biopsy when tPSA is over the cut-off and P(B) is the probability of positive biopsy for the general population. To specify the probability to have a cancer in this population the following formula could be suggested P(B)=P(B|A)*P(A)/P(A|B).

Model analysis for expected utility

A model analysis was applied to compare the screening alternatives for biochemical markers with respect to the expected utility (phase 1 of the model). Two approaches were used to solve the model. A brief explanation of both methods follows. An expected utility is the total number of cycles spent in each state (t_s), each multiplied by the incremental utility (U_s) for that state as following $\Sigma[t_s*U_s]$. This method provides a single value for each alternative, which makes it difficult for comparison. By the alternative method, Monte-Carlo simulation determines the prognoses of a large number of individual patients. Each patient begins in the starting state, and at the end of each cycle, a random number generator is used together with the transition probabilities to determine in which state the patient will begin the next cycle. When the patient enters the absorbing state, the simulation is stopped. Simulation is also stopped if the model satisfies a user defined condition (e.g. calculate utility for 5 cycles only). A *trial* refers to one individual randomly traversing a single path through the model. Individual utilities after a large number of trials form a distribution of values, from which mean and variance of expected utility can be calculated. Statistical measures (mean, standard deviation) were used to define statistical significance of the difference between alternatives, which allows to draw more precise conclusions about scenarios. These approaches for model analysis were available for the software used. A winning scenario is considered when a lower cost is associated with it.

Sensitivity analysis of the decision model

During sensitivity analysis, variables representing conditional and transitional probabilities are examined to determine how changes in assumptions affect the expected value of each strategy and the selection of an optimal strategy. In the current study the

sensitivity analysis was applied to study overall survival with respect to the changing populational characteristics (phase 2 of the model) and treatment outcome (phase 3 of the model).

The tornado diagram is an illustrative way to conduct sensitivity analysis, where results are arranged on one diagram with a single scale for all variables. On the resulting graph, each variable included in the analysis is represented by a horizontal bar. Each bar represents the range of possible outcomes generated by varying the variable within a predefined range (KEEFER et al., 2004). In the current study the horizontal bars represent varied probabilistic characteristics of the population. The range of the bars reflects the survival changes corresponding to the population changes.

Ethical considerations

Researchers did not have access to information which might be used for identification of patients. No contacts were made to individuals to whom data refer and no data linkage was performed during the raw clinical data analysis. The owners of data from Bayer clinical study on biochemical markers (mentioned earlier) gave us permission for secondary use of the data. The CHUS ethics committee gave us permission to conduct the project.

IV. Results

The "Results" chapter contains three main sections:

The first section explains model structure and how the model structure corresponds to the conceptual representation of the prostate cancer management described at the beginning of the Chapter III (structural level of model representation).

The second part introduces the results about the data used to specify the model (probabilities and utilities) with detailed review of evidence from the literature sources. Literature review was done separately for every element of the model. This section also contains the description of cut-off adjustment according to diagnostic performance of the screening program using ROC analysis (numeric level of model representation).

The last part contains the results of model analysis and sensitivity analysis aimed on identification of factors that can influence model results (functional level of model representation).

Model structure (structural and functional levels)

A study model has three phases those are screening and diagnosis, staging and treatment outcome. These phases are separately described as follows.

Model structure for Screening and diagnosis phase (phase 1)

The structure for this phase is based on the conceptual representation of prostate screening and diagnosis. Different screening scenarios were taken into account. The six alternatives described in Chapter III are shown in Table 6. These alternatives compare new to existing markers including ratios of markers and also conjunction with DRE. The traditional way of screening includes tPSA with or without DRE.

The model for the first phase contains two Markov states. The entry state represents the male population eligible for prostate cancer screening and diagnosis. The leaving state is an absorbing state representing newly diagnosed prostate cancer. The transitions between these two states can be expressed as a decision tree or a belief network. Because of the compact representation, belief networks for diagnostic tests are shown for each diagnostic strategy alternative.

Historically, the tPSA marker was introduced for screening in combination with DRE and TRUS. Later TRUS was suggested to be eliminated as a first order test. Several longitudinal studies have reported modification of screening strategies by eliminating DRE at follow up screening visits after taking into account number of cancers that could be identified by different tests alone (e.g. tPSA alone, DRE alone, etc). At the same time f/t and c/t PSA ratio have been suggested for improving specificity of the screening as an additional test to tPSA and DRE together. Recent publications on cPSA measurements in screening populations have shown that cPSA might be used instead of tPSA without decrease in diagnostic performance. These approaches for screening are represented as six scenarios for the first phase model.

Table 6. Set of diagnostic procedures for different alternative screening and diagnosis strategies

Group	Alternative	Alternative to compare
With DRE	1. tPSA+DRE+**c/tPSA**+biopsy	2. tPSA+DRE+**f/tPSA**+Biopsy
	3. **tPSA**+DRE+Biopsy	4. **cPSA**+DRE+Biopsy
Without DRE	5. **tPSA**+Biopsy	6. **cPSA**+Biopsy

Figure 7. Belief network for "tPSA+DRE+c/tPSA+Biopsy" strategy

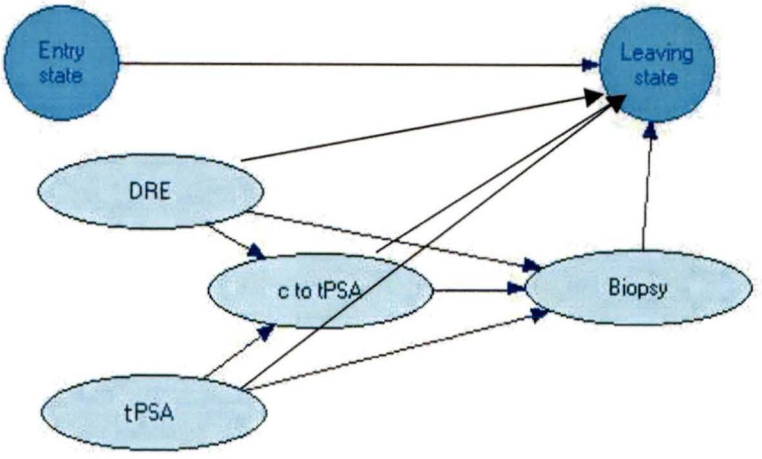
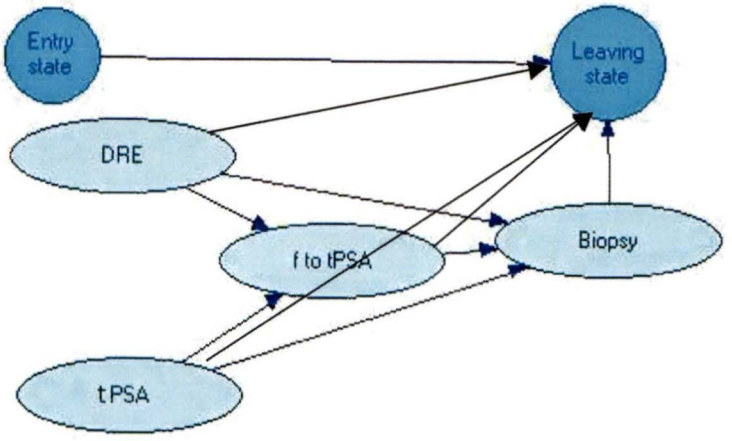


Figure 7 is a screenshot from the window in the decision modeling software ("ReasonEdge Modeler"). The "entry" state and the "leaving" state represent Markov states. The line between them is a transition between them.

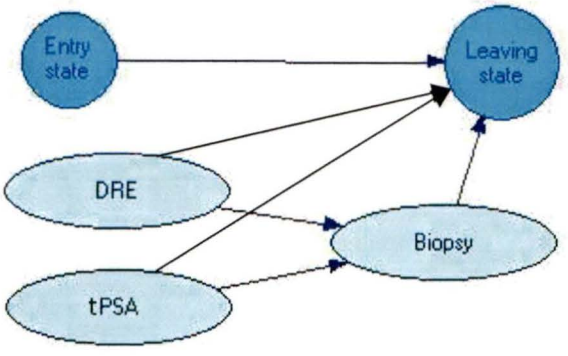
DRE and total PSA with cut off 4 μg/l are used as the first line tests. The third test (complexed-to-total PSA ratio) has been used to improve specificity. Value of total PSA is classified for three ranges (below 4 μg/l, 4-10 μg/l and more than 10 μg/l). If tPSA is more than 10 μg/l patients should definitely undergo the biopsy. All patients with tPSA 4-10 μg/l or suspicious DRE and tPSA less than 4 μg/l undergo c/t PSA ratio test. Dichotomous results of c/tPSA ratio are used to select appropriate patients for prostate biopsy using the calculated cut-off described later. A prostate biopsy is assumed to be the definitive diagnostic test for a prostate cancer.

Figure 8. Belief network for "tPSA+DRE+f/tPSA+Biopsy" strategy



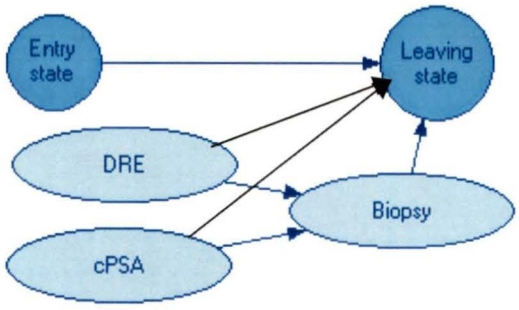
DRE and total PSA with cut off 4 μg/l are also used as the first line tests for this strategy (Figure 8). However free-to-total PSA ratio is the third test to improve specificity.

Figure 9. Belief network for "tPSA+DRE+Biopsy" strategy



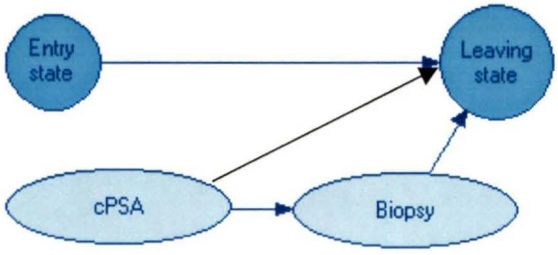
This strategy contains three tests (Figure 9). DRE and tPSA are used as the first order tests. The patient is selected for biopsy if the result of DRE is suspicious for prostate cancer and/or tPSA is above the calculated cut off limit.

Figure 10. Belief network for "cPSA+DRE +Biopsy" strategy



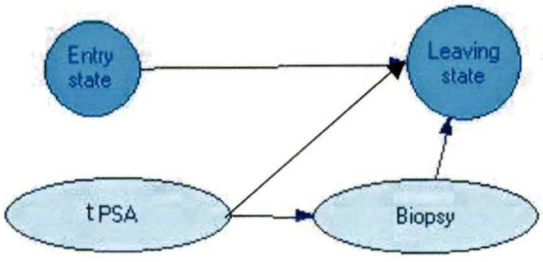
This strategy (Figure 10) is similar to the one that contains DRE, tPSA and biopsy (Figure 9).

Figure 11. Belief network for "cPSA+Biopsy" strategy



This alternative (Figure 11) represents the screening strategy where complexed PSA is used as the first order test for patients. This test is used to select patients for prostate biopsy. Complexed PSA test has dichotomous outcomes "eligible for biopsy" or "not eligible" applied using the calculated cut off.

Figure 12. Belief network for "tPSA+Biopsy" strategy

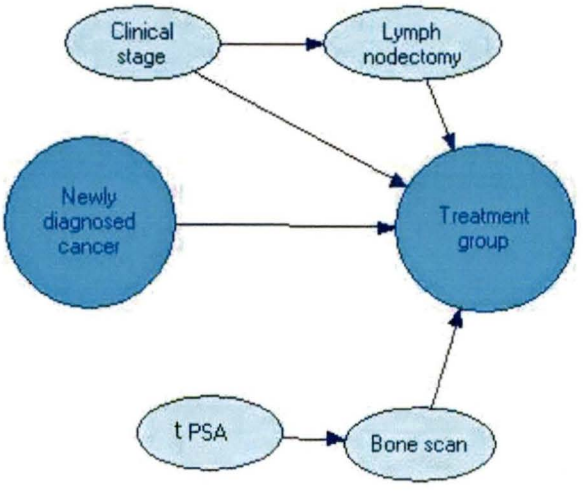


This strategy (Figure 12) is similar to the one with complexed PSA (Figure 11). cPSA is replaced by the total PSA.

Model structure for Staging phase (phase 2)

The second phase is also shown using a belief network (see Figure 13). The belief network shows the different diagnostic tests used to separate a population of men with newly diagnosed prostate cancer into different clinical groups of clinical stage. Data, used to specify the belief network, was obtained from the Bayer clinical study[11] and further combined with literature based data, which are presented in the next section.

Figure 13. Belief network for Phase 2 of the model



Two cycles ("Newly diagnosed cancer" and "Treatment group") represent beginning and end states for this phase. Other elements represent diagnostic tests used to assign patient to the appropriate treatment group. The link between the "PSA" and "Bone scan" nodes means conditional dependency between two tests (e.g. bone scan is requested if tPSA is more than 10 μg/l). Lymph nodectomy is requested if the patient is eligible for a

radical prostatectomy according to the clinical stage. Clinical stage, lymph nodectomy and bone scan results (if these tests are administered) define the treatment requirements.

The same model (Figure 13) is represented as a decision tree in Figure 15 which also shows the different treatment groups used in phase 3.

The initial assessment of clinical stage allows to separate patients into local and extended cancer groups. A tPSA value more than 10 μg/l is used to identify a high-risk population for metastases and undergo bone scan. Patients with local cancer and/or negative bone scan result are eligible for radical treatment (prostatectomy or radiotherapy). Each patient undergoes pelvic lymph nodectomy before the radical treatment is administered.

The belief network was used to facilitate calculation of conditional probabilities for newly diagnosed prostate cancer patients to be assigned to a specific treatment group on the basis of the clinical data set and literature data sources. The explanation of these calculations follows.

The distribution of patients between clinical states (by TNM classification[12]) was obtained from the clinical dataset. The probability to have a positive result on bone scan or positive lymph nodes during prostatectomy were acquired from the literature data source. The probability to have tPSA > 10 μg/l was obtained from the dataset and literature sources.

Based on the model structure for the staging phase, the probability of a positive bone scan is conditioned by the probability to have tPSA > 10 μg/l. According to the Bayesian rule, the final probability to have positive bone scan results (*Group 4*) is a multiplication of probabilities for "tPSA > 10 μg/l" and "Positive Bone scan for those who have tPSA > 10 μg/l" [**p_m1=p_psa*p_bone**].

---

[11] "Multicenter clinical performance of the Bayer Immuno1 Complexed PSA assay in the screening population" (1202 patients)

[12] See Chapter II for the details on the TNM classification of prostate cancer

The distribution of clinical stages is based on the Bayer clinical study dataset. A probability of being assigned to *Group 3* is calculated as the multiplication of probability C or D clinical stage and probability of not being assigned to the Group 4 **[p_T3*(1-p_m1)**]. All patients with clinical stage A and B are recommended to follow a radical treatment.

A probability of positive lymph nodectomy is available from literature dataset. The probability of being assigned to the *Group 2* is a multiplication of joined probabilities for p_T3 and lymph nodectomy [**(1-p_T3)*p_node**]. The probability of being assigned to the *Group 1* is a subtraction of probabilities of being assigned to other groups from 1.[13]

Table 7. A summary of information about treatment groups

Group 1	Organ confined prostate cancer; every patient should undergo lymph nodectomy before a radical treatment is assigned. If lymph nodes are positive, patient is transferred to the second treatment group;
Group 2	Pelvic lymph node positive patients
Group 3	Non organ confined cancer patients (clinical state C or D without distant metastases)
Group 4	Imaging is aimed to find bone metastases in order to avoid using prostatectomy in metastases positive patient. Patients with suspected M1 cancer (distant metastases)

The difference between the groups 2 and 3 is that C and D stage patients were assigned directly to the group 3 according to the clinical state, but group 2 contains patients with clinical state A-B and lymph node metastases. All patients with positive bone scan were assumed to be assigned to the treatment as M1 patients.

[13] List of variables used at formulas
p_psa - probability for "PSA > 10 mkg/l" among newly diagnosed prostate cancer
p_bone – probability for positive bone scan for those who have PSA > 10 mkg/l
p_m1 – probability of being assigned to the Group 4
p_T3 – probability of C or D clinical stage of prostate cancer
p_node – probability of positive lymph nodectomy for patients assigned for a prostatectomy

Model structure for Treatment phase (Phase 3)

This model contains two Markov states, "alive" and "dead". These states represent the life status of treated patients. Literature based information was used to obtain survival probability data with respect to the different treatment groups. Transitional probabilities were calculated from the overall survival rate obtained from the most recent randomized controlled trials (RCT) on evaluation of prostate cancer treatment. Details on studies and treatment is provided later in Table 19.

The patients from Group 1 are considered eligible for a radical treatment. Two main types of radical treatment are available, prostatectomy and radiotherapy. There are no published randomized trials that directly compare prostatectomy with radiotherapy in men with local prostate cancer. Radical prostatectomy has been chosen in the current project as an instance of such therapy.

The optimal treatment for men with extended disease (Groups 2 and 3) includes conservative surgery, external beam radiation with or without hormone manipulation, or hormone therapy alone. Surgery and/or radiotherapy are considered for local control, while hormone (androgen ablative) therapy is used to control a distant disease.

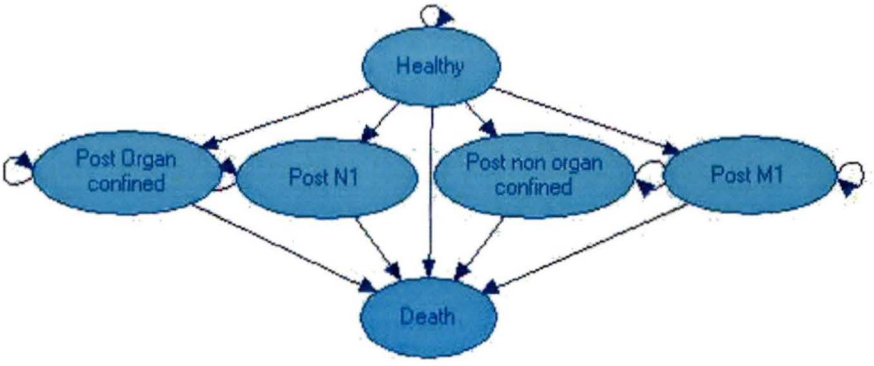
Treatment for patients with bone metastases (Group 4) is primarily palliative, and the goals are to relieve pain, improve mobility and prevent complications.

Model structure: integrating the different phases

The following six states represent the entire model. State transition representation allows compact representation in Markov model. The state transition diagram (see Figure 14) contains the following Markov states ("Healthy, or eligible for the screening", treatment outcomes, death which is an absorbing state) and allowed transitions between them. All

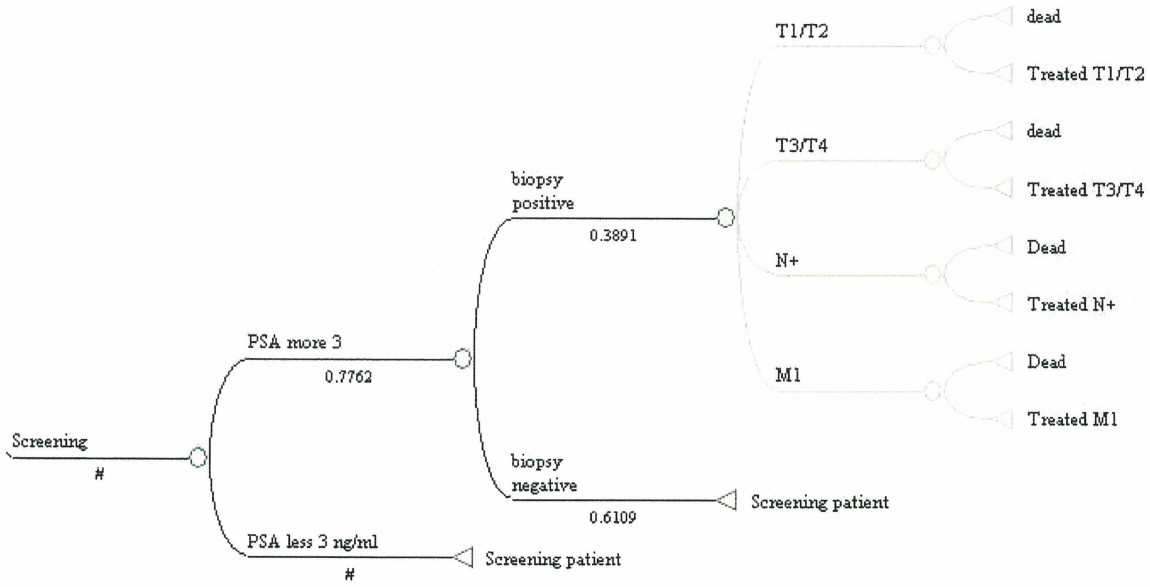
transactions from "Healthy" state to "Post ...." (post-treatment) states show the transition from newly diagnosed prostate cancer to a specific treatment group. Any transactions from these four states to themselves represent post-treatment survival. Any transaction to the "Death" state represents mortality from any reason.

Figure 14. State transition diagram for a model



Based on the software comparison (see Chapter III), Data 3.5 was used to solve the model (only decision tree representation is available). The model for Data 3.5 software was created by joining together all models representing the screening and diagnosis, staging and outcomes.

Figure 15. A fragment of decision tree represents screening and diagnosis, staging and treatment for one particular alternative (tPSA & DRE based screening)[14].



The phase 1 (up to Biopsy result on Figure 15) was obtained from the screening and diagnosis phase model (see Figures 7-12). The tree starts from the branch, which represents patients eligible for prostate cancer screening. The upper branch ("biopsy positive") represents patients with newly diagnosed prostate cancer. The staging phase 2 (four treatment groups on Figure 15) shows a part of the final model. The part 3 (treatment group survival on Figure 15) was obtained from the treatment phase model.

This table summarizes the amount of alternatives evaluated at each phase of the modeling for the prostate cancer problem.

[14] The size of the decision tree of the complete model limits our ability to provide a view on all alternatives simultaneously.

Table 8. Modeling approaches used for the three phases of building the model

	Population	Data source	Alterna-tives	Diagnostic tests
I phase (screening and diagnosis)	Referred	Clinical data set	6	tPSA, f/tPSA, c/tPSA, cPSA, DRE
	General	Published studies	1	tPSA, DRE
II phase (staging)	Newly diagnosed cancers	Clinical dataset, published studies	1	tPSA, Bone scan
III phase (treatment outcomes)	Treated cancers	Published studies	1	No tests evaluated in current model for this phase

Populations (numeric level)

Two population groups were identified for this study. The first group is based on the literature studies (referred to as "literature dataset"). This is a group of men without history of prostate cancer selected from the general population.

The second dataset are the data of the patients from the "Multicenter clinical performance of the Bayer Immuno1 Complexed PSA assay in the screening population" study (referred as "Bayer clinical study dataset"). The purpose of the study was the evaluation of new biochemical markers in a referred population.

The data about DRE, tPSA, biopsy were used to specify the screening and diagnosis phase of the model. Bone scan and lymph nodectomy results were used to specify the staging phase. Post-treatment survival rates were used to provide the transitional probabilities of the treatment phase of the model.

Literature based dataset

Target articles for DRE (screening and diagnosis phase)

Six studies were identified for populations of men eligible for prostate cancer screening, according to the selection criteria described in the methods. The populations show some heterogeneity in terms of the study population selection criteria. Catalona et al. (1994), Higashihara et al. (1996) and Crawford et al. (1996) published results for populations selected from the general community via advertisements. Brett (1998) selected patients from general practice (during routine surgery attendances). The population from the Shapiro et al. (1994) article represents patients referred by urologists to a tertiary centre. The Cooner et al. (1990) study population is a combination of referred and general practice patients. Exclusion criteria for all studies was history of prostate cancer. Catalona et al. (1994) also excluded patients with acute prostatitis or urinary tract infection. The results for DRE were normal (non-suspicious for prostate cancer) or abnormal (suspicious for prostate cancer). If the author presented results for different degrees of prostate abnormality on digital rectal examination, all these outcomes were considered as abnormal (see Table 10). The intervention (DRE) was performed by experienced examiners, the description of what was considered as abnormal DRE was clearly stated for each study.

Table 9. Summary about literature source for DRE (screening and diagnosis phase)

Publication	Population	Intervention (DRE)	Outcome
(CATALONA et al., 1994)	Responded to advertisements from general community. Exclusion: history of prostate cancer, acute prostatitis or urinary tract infection.	By urological surgeons or medical oncologist	Normal and abnormal

(BRETT, 1998)	Patients from general practice (recruited during routine surgery attendance). Exclusion: histologically confirmed prostate cancer	By author. Patients were examined lying in the left lateral position[15]	Normal and abnormal
(SHAPIRO et al., 1994)	Referred to tertiary centre by urologist from other centers (7 % from ongoing prostate cancer detection program)	By radiologist	Normal and abnormal
(COONER et al., 1990)	Patients within an urological practice (referred and general practice patients).	By 2 urologist	Normal and abnormal
(CRAWFORD et al., 1996)	Responded to advertisements from general community Exclusion: no history of prostate cancer	By experienced examiners	Normal and abnormal
(HIGASHIHARA et al., 1996)	Responded to advertisements from general community.	By urologists experienced with examining patients with prostate cancer	Suspicious and non-suspicious for prostate cancer

Table 10. Literature source dataset for DRE (screening and diagnosis phase)

Publication	DRE+	Total
(CATALONA et al., 1994).	982	6630
(BRETT, 1998)	19	211
(SHAPIRO et al., 1994)	203	471
(COONER et al., 1990)	565	1807
(CRAWFORD et al., 1996)	3107	31953
(HIGASHIHARA et al., 1996)	135	701
Total	5011	41773
Summary probability	0.12	1

The population from these studies totals 41773 patients. Heterogeneity of probabilities for DRE outcomes was evaluated using the Chi-square statistic. A heterogeneity

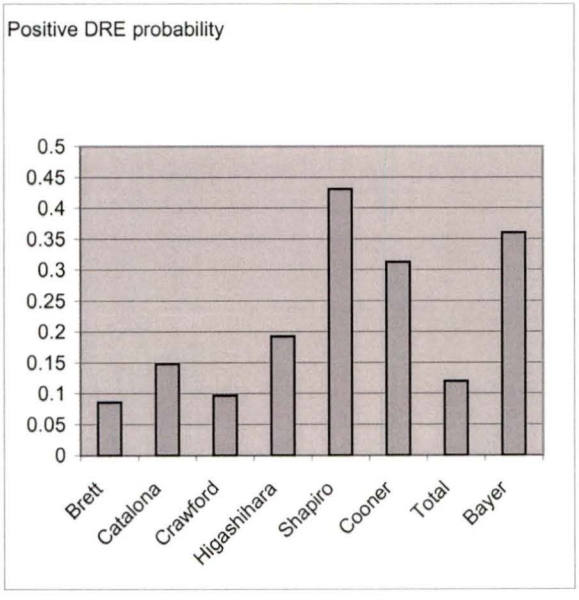
[15] No other indication about author training is available except that he works in the Urological Research Centre in Western Australia.

of results from individual studies was found to be statistically significant with $p < 0.01$. Such findings might be due to the population heterogeneity.

Two charts were generated to study the source of heterogeneity. The first one (Figure 16) shows the variability between the particular studies. The second one (Figure 17) shows the dependency between probability and study population size. Two studies where the patient population was recruited from urological practices shows that the probability of positive DRE is higher than for studies where patients were recruited from the general population (see Figure 16, studies of Cooner 1990, Shapiro 1994). The highest variability is observed for studies with smaller sample size (see Figure 17).

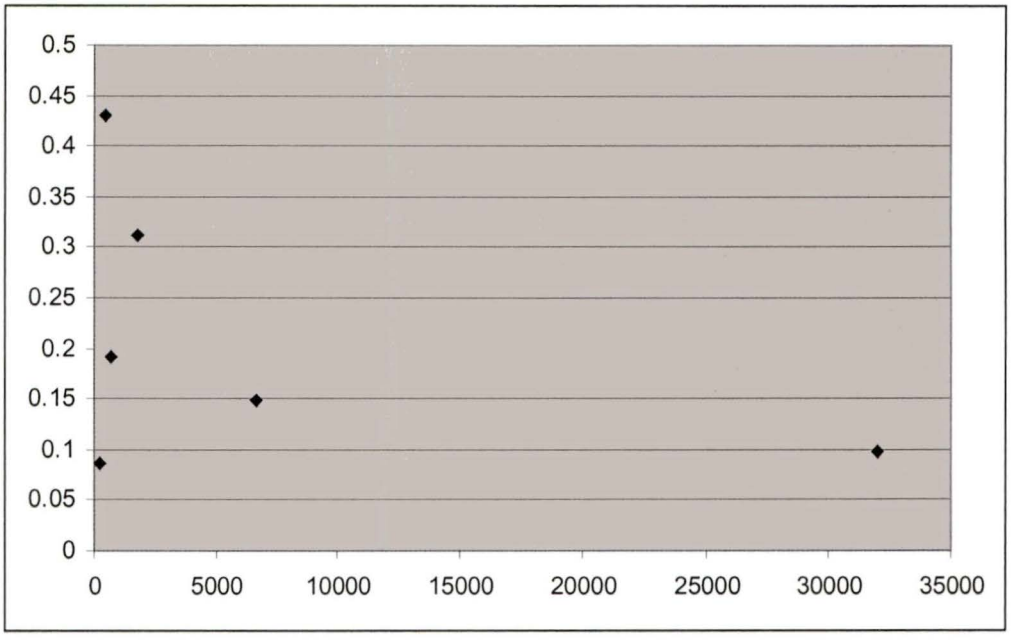
In all results on literature review a summary probability was calculated using Equation 1 as described on page 51.

Figure 16. Literature source dataset for DRE (screening and diagnosis phase)[16]



[16] "Bayer" corresponds to the data on the DRE outcomes distribution from the Bayer clinical study dataset.

Figure 17. Variability between probabilities for DRE among the studies vs. study size



Target articles for tPSA (screening and diagnosis phase)

The populations show some heterogeneity in terms of the study population selection criteria. Brawer et al. (1992), Catalona et al. (1994), Higashihara et al. (1996), Kirby et al. (1994), Thompson and Zeidman (1992), Labrie et al. (1996) selected patients from a general community via advertisements, invitation for annual check-up, etc. Other authors selected patients from the hospital departments. For example Waidelich et al. (1997) selected them from the department of internal medicine, Cooner et al. (1990) selected from urological practice. Some authors did not provide enough details on the patient population (general population or referred patients). Patients with history of a prostate cancer were not allowed to participate in the selected studies. Blood samples were obtained before DRE because any manipulation of the prostate can increase the level of tPSA.

Table 11. Summary about the literature sources for tPSA screening

Publication	Population	Intervention (tPSA)	Outcome
(HIGASHIHARA et al., 1996)	Responded to advertisements from general community.	Samples were obtained before DRE. Used the Hybrithech Tandem-R assay	0-4.0 μg/l; 4.1-9.9 μg/l; >10 μg/l
(WAIDELICH et al., 1997)	Patients from general practice (department of internal medicine) Exclusion: history of prostate cancer	Samples were obtained before DRE. Tandem-E (Hybritech)	0-4.0 μg/l; >4-10.0 μg/l; >10 μg/l
(CATALONA et al., 1994)	Men responded to advertisements from general community. Exclusion: history of prostate cancer, acute prostatitis or urinary tract infection	Blood samples were obtained before or at least 1 week after DRE	0-4 μg/l ; 4.1-10.0 μg/l ; >10 mg /l
(METTLIN, 1993)	Men (55-70 years old) Exclusion: history of prostate cancer, undergoing evaluation for prostate cancer	Blood samples were obtained before DRE	0-4.0 μg/l; >4-10.0 μg/l; >10 μg/l
(BRAWER et al., 1992)	Men (>50 years old) responded to advertisements to general community (direct mail ad; through senior registries; by lecturers at senior centers; by posters placed in hospitals) Exclusion: history of prostate carcinoma	Blood samples were taken before DRE. By Tandem-R (Hybritech, Inc)	0-4 μg/l; > 4 μg/l
(COONER et al., 1990)	Patients (50-89 years old) from within an urological practice (either were symptomatic, had been sent by a referring physician or simply worried about the possibility of cancer)	tPSA determination was delayed for >1 week after DREW or urethral manipulation. Tandem-R (Hybrithech, Inc)	0-4 μg/l ; 4.1-10.0 μg/l ; >10 mg /l
(EGAWA et al., 1995)	Males (55 years of age or older) Exclusion: no history of prostate cancer	Blood samples had been taken before DRE. tPSA IMxPSA, Abbott, Inc	0-4 μg/l; > 4 μg/l
(KIRBY et al., 1994)	Men (55 - 70 years old) from a general practice, invited by letter to attend a health check-up including prostate examination. Exclusion: no history of prostate cancer	Samples were taken before DRE. By Tandem-R Hybritech Corp.	0-4 μg/l; >4 and <10.0 μg/l; > 10.0 μg/l
(THOMPSON and ZEIDMAN, 1992)	Men (41-87 years old) invited through television and print media. Exclusion: no previous diagnosis of prostate cancer	Included data of pre-DRE tPSA assessment	0-4, 4.1-10, more 10.1 μg/l
(LABRIE et al., 1996)	Men (45 - 80 years old) were randomly selected for screening tests from the electoral rolls of Quebec city and its vicinity. Invitation had been made by the letter without any public announcement through the media.	Samples were taken before DRE and TRUS performed. Analyzed by immunoradiometric assay (Tandem-R tPSA, Hybritech)	0-4 μg/l, 4.1-10 μg/l, > 10.1 μg/l

Table 12. Literature source dataset for tPSA (screening part)

Publication	tPSA > 4 μg/l	tPSA 0-4 μg/l	N
(HIGASHIHARA et al., 1996)	79	622	701
(WAIDELICH et al., 1997)	43	219	262
(CATALONA et al., 1994)	983	5647	6630
(METTLIN, 1993)	328	1901	2229
(BRAWER et al., 1992)	187	1062	1249
(COONER et al., 1990)	602	1205	1807
(EGAWA et al., 1995)	40	1149	1189
(KIRBY et al., 1994)	103	483	586
(THOMPSON and ZEIDMAN, 1992)	327	2409	2736
(LABRIE et al., 1996)	965	7064	8029
Total	3657	21761	25418
Summary probability	0,144	0,856	

The population from these studies totals 25418 patients. Heterogeneity of probabilities for tPSA outcomes was evaluated using the Chi-square statistic. A heterogeneity of results from individual studies was found to be statistically significant with $p < 0.01$. A potential source of heterogeneity is the population selection for these studies described above. The patients recruited from urological practice have higher probability of tPSA > 4 μg/l (COONER et al., 1990). Literature source probabilities were plotted as charts to study variability (Figures 18, 19).

Figure 18. Literature source dataset for tPSA (screening part)[17]

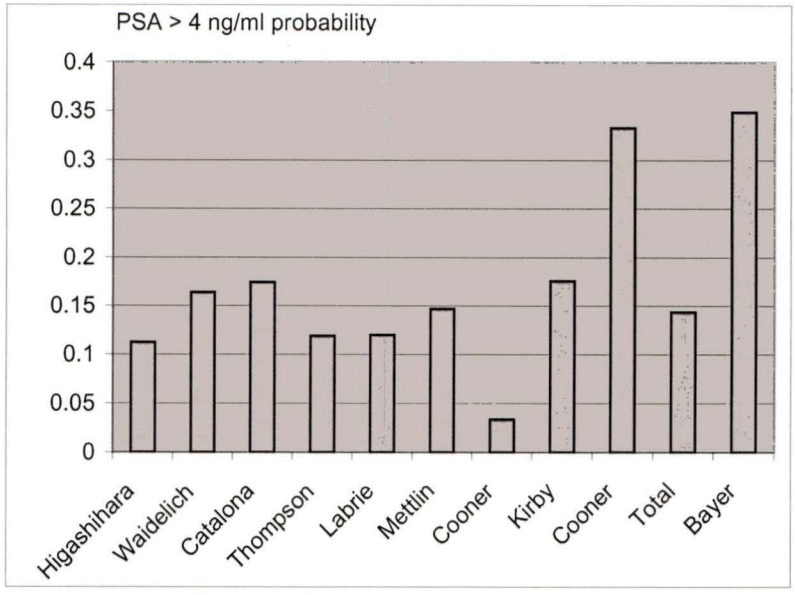
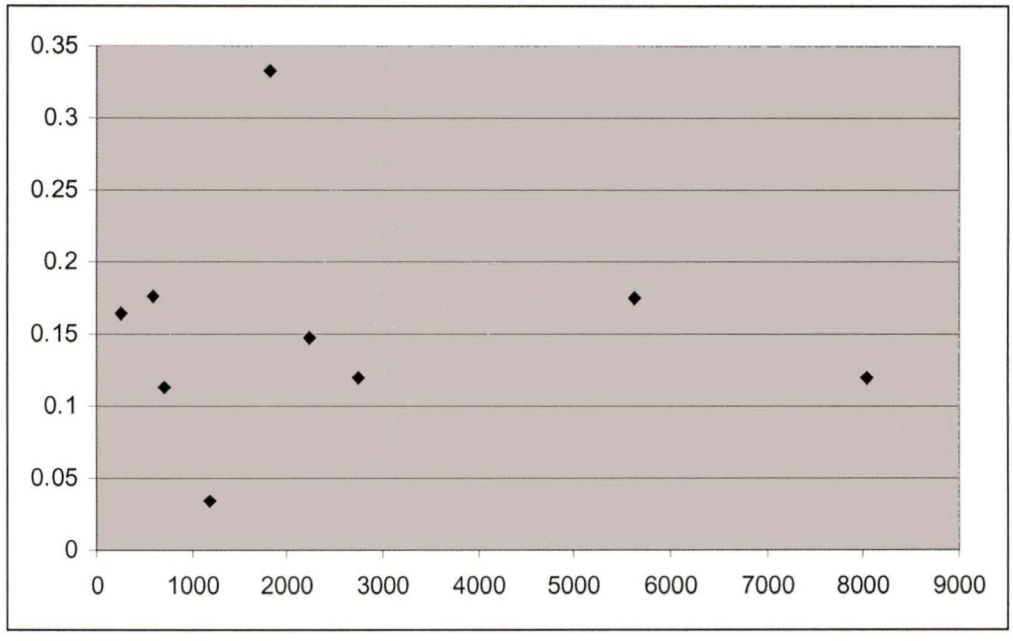


Figure 19. Variability between tPSA probability among the studies vs. study size



[17] "Bayer" corresponds to the data on the DRE outcomes distribution from the Bayer clinical study dataset.

Target articles for Biopsy (screening and diagnosis phase)

Only two studies were considered as meeting the inclusion criteria. All the patients were from the general community and had no history of prostate cancer. Biopsy was performed for all patients who had an elevated tPSA and/or DRE suspicious for prostate cancer.

Table 13. Summary of literature sources for biopsy (screening and diagnosis phase)

Publication	Population	Intervention	Outcome
(CRAWFORD et al., 1996)	Men responded to advertisements from general community Exclusion: no history of prostate	Biopsy was performed for patients who has an elevated tPSA and\or DRE suspicious of cancer	Cancer and non cancer
(CATALONA et al., 1994)	Men responded to advertisements from general community. Exclusion: history of prostate cancer, acute prostatitis or urinary tract infection	If the tPSA concentration was elevated (more than 4 µg/l) and/or DRE was suspicious for cancer the subjects underwent TRUS guided biopsy	Cancer and non cancer

Table 14. Literature source dataset for biopsy (screening and diagnosis phase)

Test 1	Test 2	Study reference	Biopsy+	Biopsy-	Total
DRE-	tPSA 0-4 µg/l	(CRAWFORD et al., 1996)	0	0	0
		(CATALONA et al., 1994)	0	0	0
		Total	0	0	0
		Summary probability	0	0	
	tPSA 4-10 µg/l	(CRAWFORD et al., 1996)	115	381	496
		(CATALONA et al., 1994)	118	366	484
		Total	233	747	980
		Summary probability	0.238	0.762	
	tPSA >10 µg/l	(CRAWFORD et al., 1996)	49	70	119
		(CATALONA et al., 1994)	35	48	83
		Total	84	118	202
		Summary probability	0.416	0.584	

DRE+	tPSA 0-4 µg/l	(CRAWFORD et al., 1996)	78	456	534
		(CATALONA et al., 1994)	48	433	481
		Total	126	889	1015
		Summary probability	0.124	0.876	
	tPSA 4-10 µg/l	(CRAWFORD et al., 1996)	129	148	277
		(CATALONA et al., 1994)	98	104	202
		Total	227	252	479
		Summary probability	0.474	0.526	
	tPSA >10 µg/l	(CRAWFORD et al., 1996)	63	120	183
		(CATALONA et al., 1994)	60	87	147
		Total	123	207	330
		Summary probability	0.373	0.627	

The population from these studies totals 3006 patients. Probabilities for biopsy outcomes were not evaluated for the heterogeneity of results from individual studies due to the small amount of studies selected.

Target articles for Bone scan (staging phase)

All patients from selected studies had newly diagnosed prostate cancer and no previous therapy initiated before the bone scan with 99mTc. The outcomes were presence or absence of bone metastases. Only patients who had tPSA more than 10 µg/l were selected to determine the extraprostatic extension of a disease. All studies where authors could not provide the information about the tPSA for patients or range of tPSA (e.g. more than 20 µg/l, 10-20 µg/l or more than 10 µg/l) were excluded from this review.

Table 15. Summary of literature sources for bone scan staging

Publication	Population	Intervention	Outcome
(CHYBOWSKI et al., 1991)	Newly diagnosed PC. No therapy was initiated before the bone scan and tPSA determination	Bone scans were performed with using 99mTc-methylene diphosphonate	Bone scan positive/negative for tPSA 0-10, >10 µg/l
(KEMP, 1995)	Newly diagnosed PC. No therapy was initiated before the bone scan and tPSA determination	All whole-body bone scan; 550 MBq technetium99m MDP; Auto-LELFIA PA kit	Bone scan positive/negative for tPSA 0-10, >10 µg/l
(RUDONI et al., 1995)	Newly diagnosed PC. No therapy was initiated before the bone scan and tPSA determination	740 MBq of 99mTc-methylene diphosphate	Positive/negative for tPSA range is 0-4; 2-10; 10-20, >20 µg/l
(GLEAVE et al., 1996)	Newly diagnosed PC. No therapy was initiated before the bone scan and tPSA determination	99mTc-methylendiphosphonate; results are evaluated by certified nuclear medicine physicians	Positive/negative for tPSA 0-10; 10-20; >20 µg/l
(HUNCHAREK and MUSCAT, 1995)	Newly diagnosed untreated PC; no prior treatment	Bone scan in local health care facility	Positive/negative for tPSA as 0-4, 4-10, 10-20, >20 µg/l
(RYDH et al., 1999)	Newly diagnosed untreated PC; no previous therapy	tPSA by AzSYM (Abbott). Bone scan with 550 MBq methylendiphospanonate	Positive/negative for tPSA 0-10, 10-20, >20 µg/l
(WYMENGA et al., 2001)	Newly diagnosed PC	700 MBq 99mTc-methylene diphosphnonate	Positive/negative for tPSA S0-10, 10-20, >20 µg/l
(ATAUS et al., 1999)	Newly diagnosed PC	20 mjuCi 99mTc-methylene diphosphonate bolus	Positive/negative for tPSA 0-10, 10-20, >20 µg/l
(LIN et al., 1999)	Newly diagnosed, untreated PC	740 MBq (20mCi) Tc-99m methylene diphospate. Two nuclear medicine physicians. If bone scan were indeterminated, additional studies (radiographs, CT or MRI, were performed to establish the final interpretation of the bone scans and to allow classification as pos/neg.	Positive/negative for tPSA 0-10, 10-20, >20 µg/l or GS 2-6, 7-10
(VIJAYAKUMAR et al., 1994)	Bone scan at the time of diagnosis of PC, prior to any treatment	20 mCi 99m-Tc-MDP, Hybritech Tandem-R immunoradiometric assay	Positive/negative for tPSA 0-10, 10-20, >20 µg/l

Table 16. Literature source dataset for bone scan staging

Positive bone scan / total patient	tPSA range (µg/l)		
	0-10	more 10	All
(CHYBOWSKI et al., 1991)	0/207	71/314	521
(KEMP, 1995)	0/10	26/88	98
(GLEAVE et al., 1996)	0/290	28/200	490
(RUDONI et al., 1995)	0/23	54/95	118
(HUNCHAREK and MUSCAT, 1995)	2/116	8/159	275
(RYDH et al., 1999)	6/111	132/325	436
(WYMENGA et al., 2001)	14/89	97/273	362
(ATAUS et al., 1999)	3/50	48/110	160
(LIN et al., 1999)	3/177	5/93	270
(VIJAYAKUMAR et al., 1994)	0/27	17/63	90
Summary probability	28/1100	486/1720	2820

The population from these studies totals 2885 patients. Heterogeneity of probabilities for bone scan outcomes was evaluated using the Chi-square statistic. A heterogeneity of results from individual studies was found to be statistically significant with $p < 0.01$ (Figures 20 and 21). A possible explanation could be the variation in the procedures and precision of the bone scan measurement.

Figure 20. Literature source dataset for positive bone scan

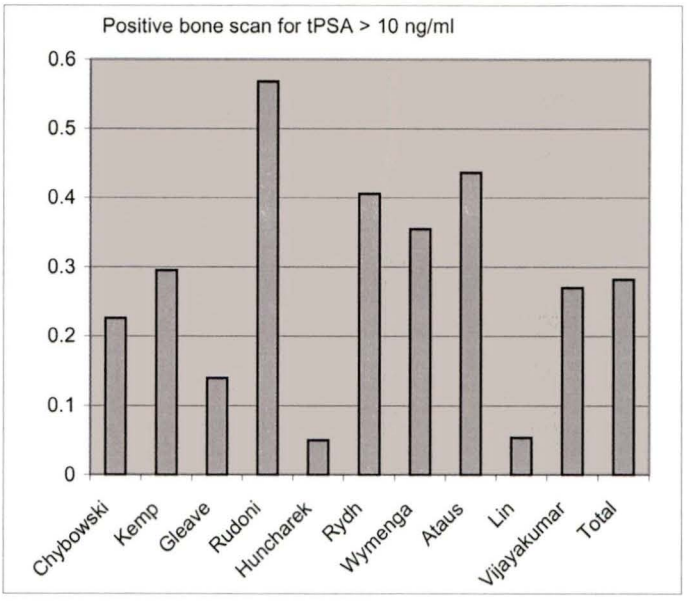
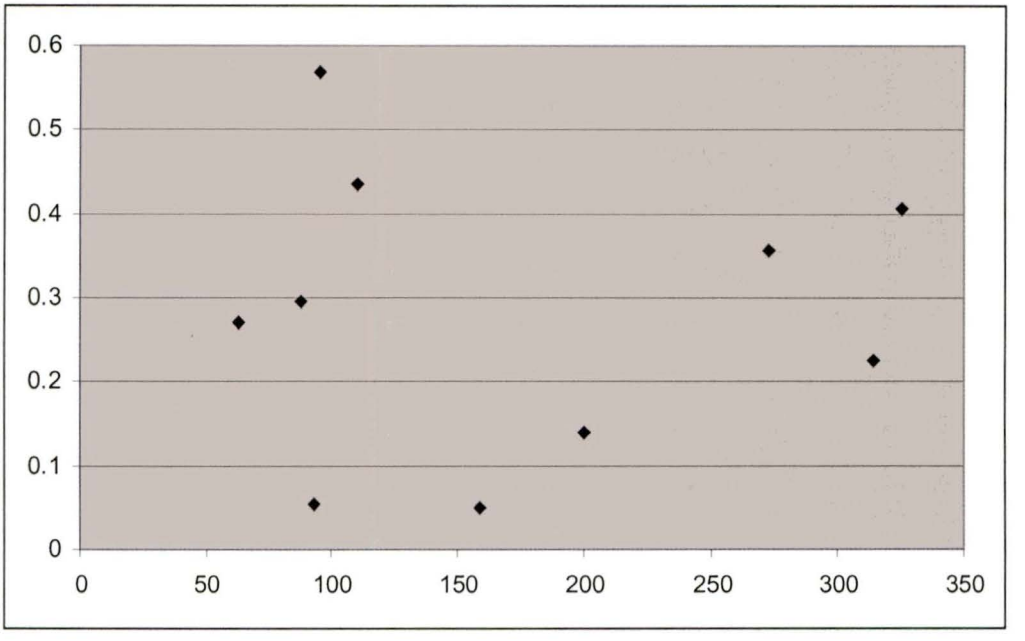


Figure 21. Variability between bone scan probabilities among the studies vs. study size



Target articles for a lymph nodectomy (staging phase)

All studies were selected to obtain the probability to have metastases in lymph nodes for patients selected for prostatectomy. Patients who had clinically local prostate cancer underwent radical retropubic prostatectomy and pelvic lymph node dissection. However the criteria for preoperative staging using imaging procedures were quite heterogeneous for these studies.

Table 17. Summary of literature sources for a lymphnodectomy

Study	N	Population	Intervention	Outcome
(ALSIKAFI, 1998)	148	Men 45 to 75 years old with clinically localized carcinoma of prostate. During this time 4 other planned radical prostatectomies were aborted when lymph node metastases were identified on frozen section; reviewer included these 4 cases	Radical retropubic prostatectomy	Lymph nodes metastases
(BADER, 2002)	236	Clinically organ confined prostate cancer; negative abdominal/pelvic CT, bone scan and chest x-ray; patients with pathologically enlarged lymph node on preoperative staging or incomplete diagnostic evaluation were excluded	Radical retropubic prostatectomy and extensive pelvic lymphadenectomy	Lymph nodes metastases
(EASTHAM and KATTAN, 2000)	475	All patients who underwent radical prostatectomy; no previous radiotherapy or hormonal therapy before radical prostatectomy.	Radical prostatectomy; laparoscopic lymphadenectomy	Lymph nodes metastases
(FERGANY, 1999)	749	Patients underwent radical prostatectomy	Radical prostatectomy; pelvic lymph nodectomy (PLND)	Lymph nodes metastases
(FREEDLAND, 2002)	815	Men underwent radical prostatectomy; some excluded due to previous chemotherapy, T0 stage; radiotherapy before operation	Radical prostatectomy	Lymph nodes metastases
(GERBER et al., 1996)	2758	Men with clinically localized prostate cancer	Radical retropubic prostatectomy; radical perineal prostatectomy	Lymph nodes metastases
(PARTIN et al., 1997)	4017	Men with clinically localized prostate cancer (T1-2)	Radical retropubic prostatectomy, staging lymphadenectomy	Lymph node metastases

Table 18. Literature source dataset for a lymphnodectomy

Study	Study size	Positive lymph nodes
(ALSIKAFI, 1998)	148	6
(BADER, 2002)	236	26
(EASTHAM and KATTAN, 2000)	475	12
(FERGANY, 1999)	749	38
(FREEDLAND, 2002)	815	16
(GERBER et al., 1996)	2758	125
(PARTIN et al., 1997)	4017	189
Total	9198	412
Summary probability	1	0.0448

The population from these studies totals 9198 patients. Heterogeneity of probabilities for lymph nodectomy outcomes was evaluated using the Chi-square statistics. A heterogeneity of results from individual studies was found to be statistically significant with $p < 0.01$. The Bader (2002) study probability is higher than others (BADER, 2002). This extreme value (see Figure 22) is smoothed by the study size weighting factor of Equation 1 in calculating the summary probability. The between studies variability is higher for studies with a smaller study population size (see Figure 23).

Figure 22. Literature source probability for a positive lymph nodectomy

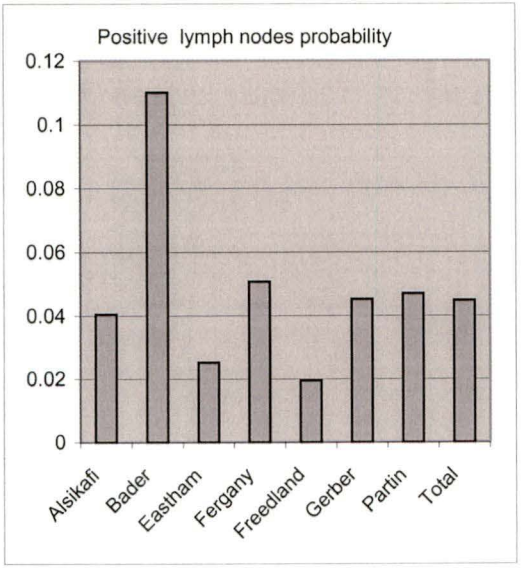
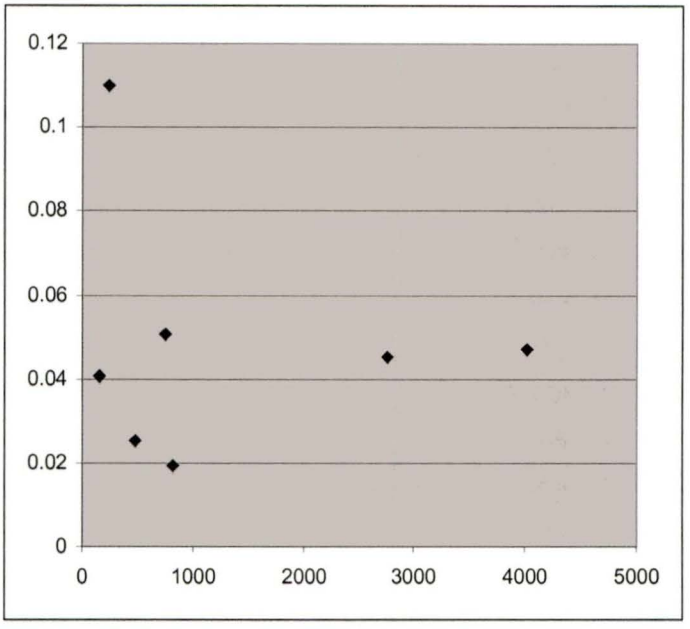


Figure 23. Variability between positive lymph nodectomy probabilities among the studies vs. study size



Target articles for post-treatment survival

Online bibliographic databases were used to select the recent clinical studies (1995-2003) of evaluation of post-treatment survival of prostate cancer patients. Keywords for the search strategy were "prostate cancer", "treatment" and "survival". The population characteristics of these studies were rigorously evaluated to match the treatment groups defined in the model. Studies were excluded if no overall mortality and/or study follow-up period were provided or made explicit. Only four studies were finally selected as suitable. An overall mortality and study follow-up period were extracted from target articles. The transformation of overall mortality to the transitional annual probability was done by the formulas described in Chapter III (page 56). Study results published by Sohayda et al. (2000) were found representative for patient survival after radical prostatectomy for local prostate cancer. Authors in this study evaluated 2424 patients treated with radical prostatectomy or radiation therapy. The treatment decision was made on the basis of patient

preferences (SOHAYDA et al., 2000). Overall survival rate in the Bolla et al. (1997) study satisfied selection needs for the group of patients with non-localized to capsule prostate cancer. Patients had to fit the following criteria before the treatment was performed, T1/T2 $N_{0\text{-}X}$[18] prostate cancer grade 3, or T3 and any grade, or T4 without involving regional lymph nodes. The clinical evaluation included bone scan, chest radiography, and CT of liver. Eligible patients had no previous treatment for prostate cancer. Two hundred and three patients from this study had been given radiotherapy and goserelin (BOLLA et al., 1997). The patients from the Burskirk et al. (2001) study represents a subpopulation of patients with positive lymph nodes (T1-4 N1 M0 stage) treated with androgen ablation plus radiation therapy (BUSKIRK et al., 2001). Glass et al. (2003) studied M1 patients' survival. Patients had histologicaly proven M1. 12.5% of them had previous prostatectomy. Patients from one of the arms underwent bilateral orchiectomy and treatment with flutamide (GLASS et al., 2003).

Table 19. Summary about literature source for treatment options

Treatment group (see Table 7)	Author	Entry state	Leaving state	Intervention	Probability of death/ 1 year
Group 1	(SOHAYDA et al., 2000)	White men with local PC	Death	Prostatectomy (or radiotherapy)	0.0187
Group 2	(BUSKIRK et al., 2001)	Pelvic lymph node positive (N1) patients	Death	Androgen ablation plus radiation therapy	0.053
Group 3	(BOLLA et al., 1997)	Locally advanced PC	Death	Radiotherapy; Radiotherapy and goserelin	0.046
Group 4	(GLASS et al., 2003)	Patients with distant metastases (M1)	Death	Bilateral orchiectomy	0.209

[18] T1/T2 stage with negative or unknown spread of disease to pelvic lymph nodes

These studies have RCT design with at least two subpopulations (control and comparison). The smallest overall mortality rate of patients was chosen among all branches for each study. The probabilities in Table 19 represents a probability to be dead during every year following treatment of prostate cancer.

Bayer clinical study dataset

A clinical study data was used to specify the probabilities for alternatives with new biochemical markers. The population are patients referred to academic and research medical centers for prostate biopsy. Strict criteria were met in order to include patients in any of the investigational sites. The following provides demographic information about the patient population, cut off adjustment technique and detailed probabilistic information about new markers according to screening scenario along with established cut offs for screening tests.

Table 20. Demographic characteristics for Bayer clinical study dataset (referred patient population)

Variables	Subjects	Mean	SD
Total PSA (μg/l)	1202	7.49	23.53
Complexed PSA (μg/l)	1202	6.38	20.98
Complex-to-total PSA ratio	1202	0.82	0.1
Free PSA (μg/l)	1202	1.06	2.47
Free-to-total PSA ratio	1202	0.15	0.07
Age, years	1201	63.28	8.75
tPSA range	0-4 μg/l	4-10 μg/l	>10 μg/l
	420	603	179
DRE	Normal	Abnormal	
	769	433	
Clinical stage	A or B	C or D	
	386	16	
tPSA range for newly diagnosed prostate cancer	tPSA < 10 μg/l	tPSA < 10 μg/l	
	317	87	

The clinical dataset contained 1202 patients. These patients were used to specify probabilistic information for screening alternatives. Based on biopsy result, 402 prostate cancer cases were identified. Clinical stages were available for all newly diagnosed cancer cases as well as a preoperative total PSA level.

When multiple markers are used in a sequence in diagnostic program, a small change in diagnostic performance of the first marker influences the use of all consecutive markers which will result in overall resource utilization changes. Based on the alternatives previously identified the cut offs for each marker were calculated. The details are provided in the following section.

Cut-offs adjustment

For multiple tests diagnostic programs, using tests in a sequence may depend on the results of previous tests. Cost of a multi-test diagnostic program is the cumulative value of cost of every test multiplied by the number of times every test has been used. Cost of different alternative programs may simply be different due to different diagnostic performance. In our example, the number of biopsies performed depends on the number of patients selected by biochemical marker(s) and/or DRE. Targeting very high or very low sensitivity is not a goal for prostate cancer screening. For example, selecting 100% sensitivity of diagnostic program results in selecting every patient for a biopsy without taking into account tPSA result. Selecting 5-10% sensitivity may result in high specificity, but biopsy may not be necessary to confirm prostate cancer. Based on these arguments, the cut offs for diagnostic tests were adjusted to support the same sensitivity level for all screening

programs. A sensitivity of 90% was chosen. The same level was calculated for all alternatives.[19]

Multiple tests do not allow simple use of ROC curve analysis (METZ, 1978). An indirect way was used to visually choose the appropriate cut off. If a diagnostic strategy consists of more than one test, the last one was plotted with output as a continuous variable on the ROC curve. A point at the ROC curve for this particular test was estimated according to the previous tests. E.g. for tPSA and DRE based strategies tPSA cut off was estimated after a certain group of patients was selected by positive DRE. So the tPSA cut off is lower than it would be if no DRE test was used before. The calculations follows.

Table 21. List of variables used for calculations (DRE + New_Test strategy)

Biopsy	Biopsy positive	Biopsy negative
DRE+	a_1	a_2
DRE- & Test+	b_1	b_2
DRE- & Test-	c_1	c_2

The three lines in the table represent three groups of patients according to biopsy status by the results of DRE and biochemical test. These are positive DRE (a_1 and a_2), negative DRE and a positive test (b_1 and b_2), negative DRE and a negative test (c_1 and c_2) vs. positive/negative biopsy results. Sensitivity for DRE and a biochemical test could be calculated as $S=(a_1+b_1)/(a_1+b_1+c_1)$. As said before, this sensitivity has been chosen as 0.9. Results of a biochemical test for patients with negative DRE could be used to plot a ROC curve for biopsy results. The sensitivity of a biochemical test for these patients could be calculated as $S^*= b_1/(b_1+c_1)$. Combining the two formulas the next equation is obtained $S^*= S - (a_1)/(b_1+c_1)$, where a_1, b_1+c_1 are known. So $S^*=0.8164$ is the cut off sensitivity for a

[19] The traditional strategy of PSA> 4 mkg/l and/or positive DRE when applied to the patient population from Bayer clinical study ("referred patients" population) supports 91-92% sensitivity.

biochemical test for patients with negative DRE to support 0.9 sensitivity of overall screening program. With known S* the appropriate cut off for a biochemical test could be chosen from the ROC curve.

Table 22. List of variables used for calculations ((DRE and/or tPSA>4μg/l) + New_Test strategy)

Biopsy	Biopsy positive	Biopsy negative
[(DRE+, tPSA <10 μg/l) or (DRE-, tPSA 4-10 μg/l)] & NewTest+	a_1	a_2
[(DRE+, tPSA <10 μg/l) or (DRE-, tPSA 4-10 μg/l)] & NewTest-	b_1	b_2
tPSA > 10 μg/l	c_1	c_2
DRE-, tPSA <4 μg/l	d_1	d_2

Four groups were identified according to the two possible outcomes for DRE results (positive or negative), three outcomes for tPSA test (less than 4 μg/l, 4-10 μg/l, more than 10 μg/l) and two outcomes for another biochemical test (positive or negative by the specific cut off to be identified). The sensitivity of the overall program is $S=0.9=(a_1+c_1)/(a_1+b_1+c_1+d_1)$, and the sensitivity of a new biochemical test is $S^*= a_1/(a_1+b_1)$ for the patients who are selected by the result of DRE and tPSA test. So $S^*=S - (0.9*d_1-0.1*c_1)/(a1+b1)$, where $S=0.9$, d_1, c_1, a_1+b_1 are known. With known S* appropriate cut off for a biochemical test could be chosen from the ROC curve as was done for the previous example.

Based on the calculations detailed in the previous paragraph, the following cut offs were obtained for different alternatives.

Table 23. Adjusted cut offs for biochemical test for alternative strategies[20]

Strategy	Adjusted cut offs (90 % sensitivity)
1. tPSA, DRE, c/tPSA based strategy	c/tPSA= 0.657 μg/l
2. tPSA, DRE, f/tPSA based strategy	f/tPSA= 0.245 μg/l

[20] Number of an alternative scenario corresponds to the number of the same alternative in Table 6 at the beginning of Chapter III.

3. tPSA+DRE based strategy	tPSA=4.17 µg/l
4. cPSA+DRE based strategy	cPSA=3.52 µg/l
5. tPSA based strategy	tPSA= 2.97 µg/l
6. cPSA based strategy	cPSA=2.46 µg/l

The following table was used to specify conditional tables for chance nodes for every alternative.

Table 24. Probabilistic information from Bayer clinical study dataset

Alternative strategy in the model	Diagnostic test	Probabilities associated with outcomes
1. tPSA, DRE, c/tPSA	tPSA	0-4 : 0.3475; 4-10 : 0.5029; more 10 : 0.1495
	DRE for tPSA 0-4 µg/l	positive: 0.4303; negative: 0.5697
	c/t PSA for DRE +, tPSA 0-4 µg/l	1: 0.9162; 0 : 0.0838
	biopsy for tPSA more 10 µg/l:	positive : 0.4804; negative: 0.5196
	biopsy for c/tPSA=1, DRE+, tPSA 0-4 µg/l	positive : 0.2805; negative: 0.7195
	c/tPSA for tPSA 4-10 µg/l	1: 0.9767; 0 : 0.0233
	biopsy for CT=1, tPSA 4-10 µg/l	positive : 0.3929; negative: 0.6071
2. tPSA, DRE, f/tPSA	tPSA	0-4 : 0.2957; 4-10 : 0.5196; > 10 : 0.1846
	DRE for tPSA 0-4 µg/l	positive: 0.4407; negative: 0.5593
	f/tPSA for DRE +, tPSA 0-4 µg/l	1 : 0.7628; 0: 0.2372
	biopsy for tPSA more 10 µg/l	positive : 0.4434; negative: 0.5566
	biopsy for f/tPSA =1, DRE+, tPSA 0-4 µg/l	positive : 0.2521 ; negative :0.7479
	f/tPSA for tPSA 4-10 µg/l	1 : 0.9437; 0 : 0.0563
	biopsy for f/tPSA =1, tPSA 4-10 µg/l	positive : 0.3986; negative: 0.6014
3. tPSA and DRE	tPSA	1:0.6294; 0 : 0.3706
	DRE for tPSA=0	positive: 0.4302; negative: 0.5698
	biopsy for tPSA=1	positive: 0.4111; negative: 0.5889
	biopsy for tPSA=0 and DRE pos	positive: 0.2723; negative: 0.7277
4. cPSA with DRE	cPSA	1 : 0.5935; 0 : 0.4065
	DRE for cPSA=0	positive : 0.4148; negative : 0.5852
	Biopsy for cPSA=1	positive : 0.4318; negative: 0.5682
	Biopsy for cPSA=0 and DRE positive	positive: 0.2772; negative: 0.7228
5. tPSA without DRE	tPSA	1 : 0.7762 ; 0 : 0.2238
	Biopsy for tPSA=1	positive : 0.3891; negative : 0.6109
6. cPSA, no DRE	cPSA	1 : 0.7554; 0 : 0.2446
	biopsy for cPSA=1	positive: 0.3998; negative: 0.6002
Stage determination	Clinical stage distribution	stage A or B : 0.96; stage C or D : 0.04
	tPSA distribution for newly diagnosed prostate cancer	(tPSA < 10 µg/l) : 0.785 ; (tPSA > 10 µg/l) : 0.215

The cost estimation for the current study is based on the cost of diagnostic procedures and health care professional rate per patient. Cost of biopsy evaluation is

calculated from pathologist rate, time for specimen preparation and cost of materials. Each biopsy includes cost for trans-rectal echography performed by a radiologist. Abdominal echography is always done before the prospective biopsy. Cost of biochemical marker is chosen to be the same for all of them (tPSA, cPSA, fPSA). Cost of ratio is calculated from two tests (costs of total PSA and complexed (free) PSA). The details are provided in Table 25.

The cost per cancer detected is also provided for screening alternatives.

Table 25 Details on costs for diagnostic procedures

Category	Cost (CAD)
Pathologist	50.00
Time and materials	54.00
Additional specimen treatment 15.00$ (1/6 patients)	If necessary
Additional specimen evaluation 18.00$ (1/6 patients)	If necessary
Total biopsy cost for 1 standard case	104.00
Radiologist	42.75
Abdominal echography (always before the biopsy)	30.00
Trans-rectal echography	45.00
Total TRUS cost for 1 standard case	117.75
Marker analysis, time and materials	7.00
Total biochemical marker cost for 1 test PSA test	7.00

Model analysis

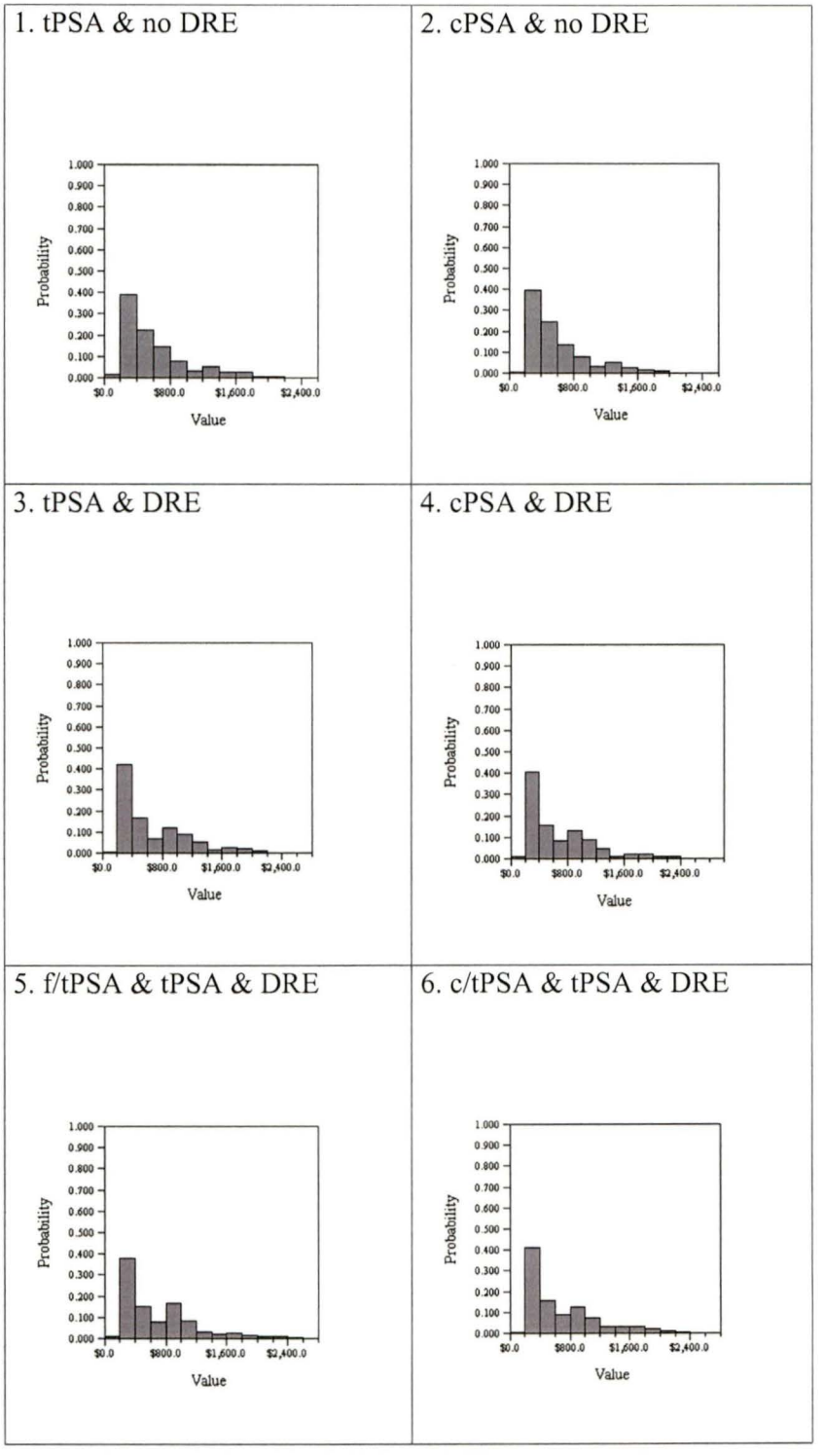
The final model contains six alternatives in the screening and diagnosis phase. These alternatives are based on the referred patients dataset. An additional alternative is based on DRE and tPSA distributions from the published studies of a general prostate cancer

screening population. Characteristics (clinical and pathological) of newly diagnosed prostate cancer patients identified by screening was assumed equal for all alternatives. This assumption allows to apply the staging phase of the model to both referred and general population of patients. Overall survival rates were applied in the treatment groups in the treatment phase. Survival rates were transformed into transition probabilities. Utilities were attached to the branches of the decision tree. Solving the model allowed to obtain utilities for expected costs, number of cancers detected and to estimate an overall population mortality.

Costs estimation by 5 year simulation analysis

Two ways of expected utilities calculation were used (rollback and Monte-Carlo simulation). Using the first method terminal branches are folded back by calculating an expected value for each terminal node. The model created in Data 3.5 was executed for 5 years of annual screening. The result is a single value for every branch deviating from the decision node. Using the second method randomly chosen outcomes were generated for every branch during a predefined number of cycles (1000 cycles were used for current model). The result is a distribution of outcomes for every branch. So results were estimated in term of distribution parameters (e.g. mean and standard deviation). The simulation data are shown as histograms for six alternatives representing different screening strategies applied to the population of referred patients. The horizontal axis represents the range for expected utilities (cost of screening during the first five years).

Figure 24. Simulation data histograms for the cost of screening alternatives (Monte-Carlo simulation, 1000 cycles)



The calculated costs include only the costs for diagnostic procedures used for a screening. The projected utility of the model does not include the expenses due to staging and treatment options. Histograms for all six simulation datasets clearly shows skewness of distribution. This confirms findings of Briggs and Sculfer (1998), cited in Chapter 1.

Table 26. Expected utility for prostate cancer model

N	Alternative strategies	Costs (Rollback method)	Costs, mean (Monte-Carlo simulation)	Costs, SD (Monte-Carlo simulation)	Cancers, mean (Monte-Carlo simulation)	Cancers, SD (Monte-Carlo simulation)	Cost/ cancer detected
1	tPSA & no DRE	553	545.8	318.95	0.820	0.384	676
2	cPSA & no DRE	539	535.99	302.28	0.829	0.376	659
3	tPSA & DRE	653	641.11	372.61	0.835	0.371	798
4	cPSA & DRE	633	630.08	360.13	0.840	0.366	773
5	c/tPSA & tPSA & DRE	657	652.00	375.43	0.803	0.397	802
6	f/tPSA & tPSA & DRE	655	655.38	379.7	0.823	0.381	800
7	DRE & tPSA literature	490	n/a	n/a	0.205	0.403	2201

The third column in the Table 26 for a 5 years simulation analysis represents values obtained from the rollback calculation of the expected costs for a 5 years time horizon. The others columns contain results of Monte-Carlo simulation over 1000 cycles. Means and standard deviations were used for expected utilities comparisons for the different pairs of diagnostic strategies (see Table 27).

Several pairs of alternative strategies were compared using Student *t*-test. Two types of comparison were performed. The first type contains tests with same number of screening tests, but different biochemical markers (e.g. tPSA vs. cPSA, f/tPSA vs. c/t PSA).

The second type compares the influence of DRE on screening strategy cost estimation.

Table 27 Alternative strategies comparison

Screening strategy, mean (SD)		**Student *t* test**
similar strategies, different markers		
tPSA & no DRE $545.8 ($318.95)	cPSA & no DRE $535.99 ($302.28)	p= 0.501
tPSA & DRE $641.11($372.61)	cPSA & DRE $630.08 ($360.13)	p= 0.841
f/tPSA & tPSA & DRE $655.38 ($379.7)	c/tPSA & tPSA & DRE $652.00 ($375.43)	p= 0.481
protocols including same markers with DRE and without DRE		
tPSA & no DRE $545.8 ($318.95)	tPSA & DRE $641.11($372.61)	$p < 0.001$
cPSA & no DRE $535.99 ($302.28)	cPSA & DRE $630.08 ($360.13)	$p < 0.001$

There was no significant difference with $p > 0.05$ between the strategies, which includes alternative markers (e.g. using tPSA vs. cPSA). The simulation datasets were also used to show the difference ($p < 0.001$) of using DRE for screening strategies.

The last columns in Table 26 represent the costs per cancer detected, which is used for efficacy evaluation of prostate cancer screening programs. As was stated before, the main complaint about the current screening programs is its low specificity. So cost per cancer detected allows to estimate the "cost" of specificity for different programs. Cost of diagnostic procedures for prostate cancer screening were estimated for 5 years. No confidence intervals were provided for the ratio due to complexity of calculations.

Last line in Table 26 represents a tPSA + DRE screening strategy based solely on literature data. No diagnostic sensitivity is available for this screening program. Hence we cannot compare it with other alternative strategies. However almost 2.5 fold changes in cost per cancer detected show a significant difference between referred patients and general populations, associated with the lower prevalence in the general population.

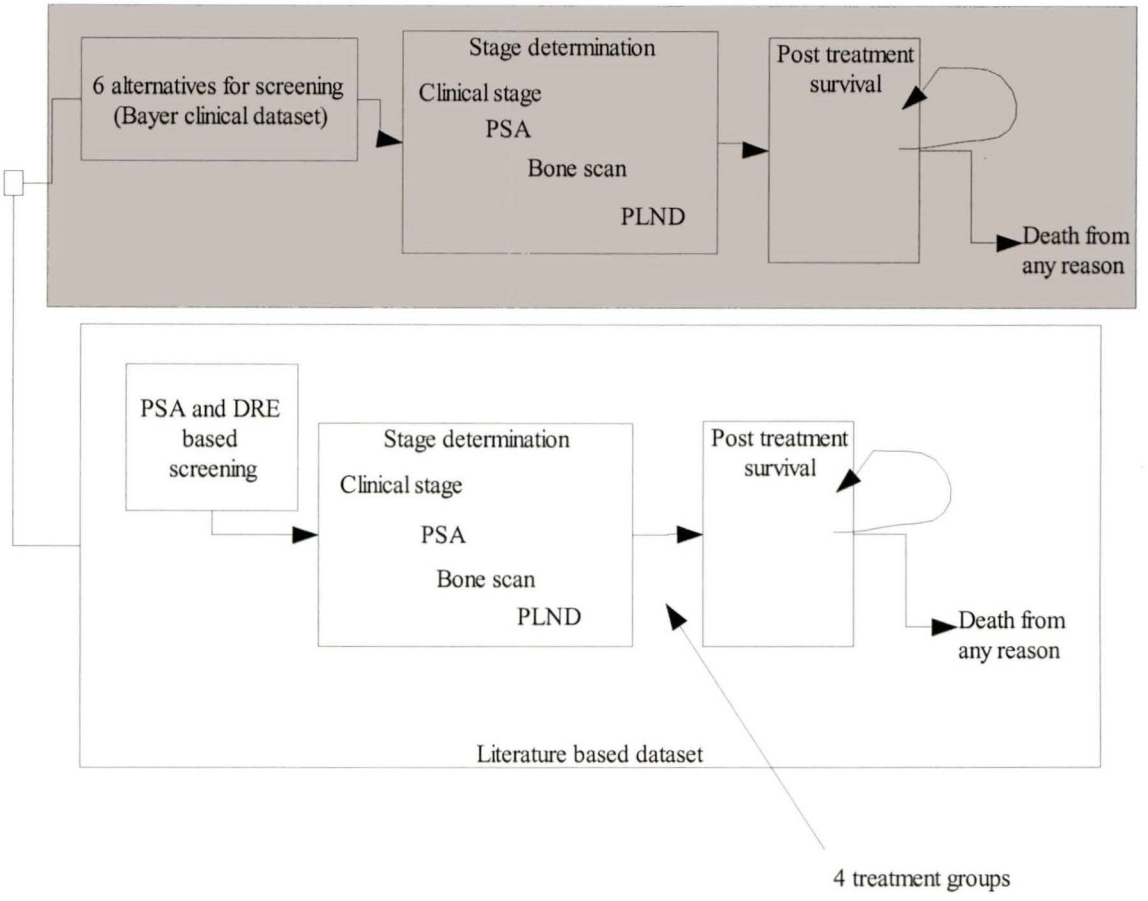
Based on simulation results there is a difference in projected 5 year costs between strategies which include DRE vs. those which do not include DRE. In other words for the same level of sensitivity, strategies without DRE found to be less expensive for serial screening.

These results respond to the first part of the second objective to evaluate scenarios of using new markers for prostate cancer screening. In order to evaluate the second part (studying model behavior) a model simulation was undertaken for expected overall mortality of a population.

Tornado diagram as a representation of decision model's sensitivity analysis

A sensitivity analysis was used to estimate how overall survival of patients with newly diagnosed prostate cancer depends on changes in treatment advancement and changes in population. Sensitivity analysis requires systematic examination of all variables implicated in the decision model.

Figure 25. Part of the model used for sensitivity analysis

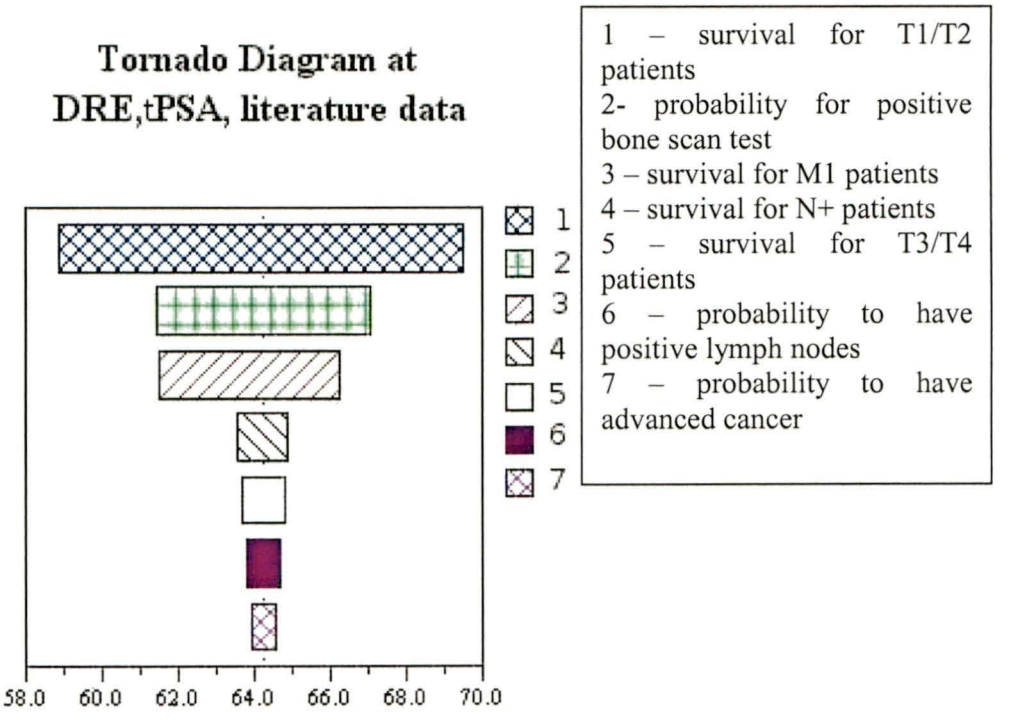


In order to apply sensitivity analysis, an alternative representing screening of the general population was created. A screening and diagnosis phase population was specified based on the literature data on newly diagnosed prostate cancer cases in the population undergoing tPSA and DRE based screening. This type of screening (tPSA cut off 4 mg/l and/or positive DRE) is commonly accepted in North America. It was the only type of screening with data easily available from the literature source. This source provided information on the rate of new cancer presentation.

To visualize several variables the Tornado diagram representation approach was used based on 5 years overall survival rate. The distribution to one of the treatment groups is determined by application of the staging phase Bayesian model (Figure 13). The

probabilistic parameters for treatment group survival and distribution of patients between treatment groups were varied within +/- 50% range[21]. The resulting changes in overall survival of the screened population were arranged on the one diagram as bars with a common scale for all variables. The scale on Figure 26 represents number of deaths per 1000 people during a 5 year annual screening.

Figure 26. Sensitivity analysis for probabilistic parameters in order to estimate influence on overall survival (represented as a Tornado diagram)



The size of the bars represents the range of possible outcomes. Variables are automatically ordered by the size of bars. In the current study survival rate for organ-confined prostate cancer (T1/T2), number of M1 cases of disease and overall survival rate for

[21] The limited number of studies selected for probabilistic parameters used for sensitivity analysis do not allow to define a more precise range for every parameter. See Chapter V for more information.

M1 stage of prostate cancer were found to produce a larger impact on the estimation of overall survival for prostate cancer patient population than the other parameters.

Summary of findings

Methodology for model creation

The following methods were used in the current project: decision analysis, systematic review and aggregation, statistical analysis. The decision analysis approach was used to study the problem. The result of this approach was a decision model, based on the conceptual representation of prostate cancer screening, diagnostic and management. Based on the model structure the data were selected using a systematic review approach. During the systematic review of scientific publications, multiple instances of evidence were identified for the decision model. An aggregation method was used to aggregate similar evidence in order to increase the robustness of the model results. Based on the model simulation results, statistical analysis was used to study the research objectives (choosing the optimal screening strategy and evaluating the new screening marker).

Model structure

A study model was created in three phases. At the first phase, the structure of the screening part was identified. Two populations were studied (referred patients and general population of men) due to available data. Six scenarios were created for a referred patients population. They include tPSA and cPSA based alternative screening scenario with/without DRE, using c/tPSA and f/tPSA ratio was also studied. For the general population, a common approach for screening with DRE and tPSA > 4 μg/l criteria was studied.

At the second phase, model structure represents treatment groups determination. The structure of the second phase model uses 4 parameters: clinical stage, tPSA cut off, bone

scan and pelvic lymph nodectomy for treatment group determination. Based on the treatment groups, the post-treatment survival of the patients was modeled during the third phase of the model. Then these three parts of the model were joined and the final model was created.

Populations

Two populations were studied (referred patients and general population). New marker scenario were studied using referred patients data available from recent clinical study. The general population data from literature source contributed to the following parts of the model (pre-treatment state determination and post-treatment survival). The screening scenario using literature data was created for studying question of how survival reflects populational changes.

New marker scenario supports the same level of diagnostic sensitivity. Cut-off adjustments were necessary to avoid a bias while comparing cost and efficacy parameters between alternatives.

Model analysis

The results of model simulation and sensitivity analysis are summarized in Table 26. From the cost comparison based on simulated results, the following conclusions are obtained. The first is that either tPSA or cPSA can be used in the screening protocol, as long as the same level of diagnostic sensitivity is maintained, there would not be a significant difference in costs for tests prescribed. The second is that the presence of DRE in the screening strategy may significantly increase the cost of screening programs over a 5 years interval. The cost per cancer detected as a parameter of cost/efficacy evaluation, has shown the difference between the programs which include DRE and those, which do not include DRE. The significant difference between the results for two populations (referred patients vs.

general population) allows a decision maker to justify the health care resources routed for prostate cancer detection.

Sensitivity analysis of variation in probabilistic parameters of treatment group distribution and group survival finds greatest variation in overall survival in the T1/T2 treatment group. This suggests that investment in detection and disease management strategies for T1/T2 survival is likely to have more impact than for the other treatment groups.

Table 28. Summary about all findings from the model

Analysis	Utility	Findings
Utility comparison	Costs	❑ No significant difference between using tPSA or cPSA in similar strategies (e.g. tPSA vs. cPSA without DRE; tPSA & DRE vs. cPSA & DRE) ❑ Significant difference between screening programs with DRE vs. without DRE
	Cost/ Efficacy	❑ Difference in cost per cancer detected for programs with DRE and without DRE (without testing the hypothesis on statistical significance) ❑ Difference in cost per cancer detected between two populations (referred patients vs. general population)
Sensitivity analysis	Overall survival	❑ All diagnostic screening strategies give similar results in a sensitivity analysis for overall mortality ❑ Mortality mainly depends on the estimation of the annual probability of death for T1/T2 and M1 treatment groups.

The sensitivity analysis was applied to the overall survival of prostate cancer patients. When probabilistic parameters were changed all strategies gave similar results. The tornado diagram helps visualize the impact of these parameters and allows to compare their impact. Overall survival mainly depends on the estimation of the annual probability of death for T1/T2 and M1 treatment groups.

V. Discussion and conclusions

Study rationale

Biochemical markers play the most important role in prostate cancer detection and management. Since 15 years they have been used in clinical practice for prostate cancer screening. Markers aim to improve "DRE only" based detection of cancer cases in earlier stages. However biochemical marker based screening raises several questions including whether screening is beneficial or not.

The tPSA marker was first introduced to clinical practice in the early 1990's as a screening test complementary to DRE. Various screening protocols have been suggested during the years of using tPSA in clinical practice. tPSA based screening, using the conventional cut-off of 4 μg/l can provide a high level of sensitivity (up to 90-95%), but the specificity remains low (20-35%), which leads to unnecessary biopsies. Later tPSA became the primary test for prostate cancer screening because the majority of cancers were associated with negative DRE. During serial screening the ability of DRE itself to identify new cancers further decreases. Some researchers have suggested that DRE might be omitted on follow up screening.

The derivates (e.g. fPSA, cPSA, f/tPSA ratio, c/tPSA ratio) were subsequently introduced as second line tests to improve specificity of prostate cancer screening. With these second line tests up to 23 % of biopsies can be avoided for certain groups of patients (e.g. patients with tPSA 4-10 μg/l) (OZDAL et al., 2004). Several recent studies have shown that based on diagnostic performance results cPSA may serve as an alternative to tPSA. The set of alternatives evaluated in the current model respond to these screening scenarios.

Several researchers have described the demographic changes in prostate cancer incidence linked to the implementation of screening. This has raised many questions whether such changes might be due to screening or due to other reasons. Conclusive determination of the beneficial effect of screening on a population requires a long follow-up period and large study size and such studies are currently underway (ERSPC[22] and PLCO[23]). A modeling approach combining results from several studies has been an approach to partially circumvent the need for major and costly longitudinal studies.

Starting in 1992, various authors have used a modeling approach to compare various modifications to the screening strategy as well as to evaluate the impact of screening on life expectancy and quality of life. Decision analysis modeling avoids some data limitations in clinical research. There is no need to have a single source dataset with all necessary variables for the analysis. Several sources of evidence can be used for specifying the decision model, from scientific publications and high quality data from clinical practice. A discussion about using decision models for evidence integration is presented later in this chapter.

Modeling process

Prior to development of a decision model, a conceptual model has to be developed. The conceptual model represents the knowledge about the problem domain. The conceptual model limits representation to the problem identified by the study objectives. Concepts and the relationships between them are used to identify the major elements of the decision model and to establish its structure. Two clinical experts, a urologist and a radiologist) were consulted to validate the conceptual model.

---

[22] European Study of Screening for Prostate Cancer (ERSPC). See http://www.erspc.org/ . Accessed on the 19th of July 2004.

[23] Prostate, Lung, Colorectal and Ovarian Cancer Screening Trial (PLCO). See http://www3.cancer.gov/prevention/plco . Accessed on the 19th of July 2004.

Conceptual models of the same clinical problem can differ in perspective as shown in the following two examples. In a study of the natural history of prostate cancer, Ross et al. (2000) focused on the successive phases of cancer progression (local and distant) and used these phases as states in a Markov model (ROSS et al., 2000). In a study of the clinical and economical burdens of prostate cancer, Grover et al. (2000a) focused on identifying the phases of prostate cancer management (diagnostics and treatment) and represented these as elements of the decision model (GROVER et al., 2000a).

Two published evaluation frameworks were considered in this study and combined to establish an enhanced set of criteria to facilitate accurate model building. Cuzin et al. (1998) created an evaluation framework for assessment of prostate cancer decision models (CUZIN et al., 1998). Drummond and Jefferson (1996) published an evaluation framework for quality of content (referred to here as the BMJ framework) in order to guide authors before submission of economic assessment articles to the BMJ (DRUMMOND and JEFFERSON, 1996). These main groups of criteria were identified as useful for model evaluation: aspects of modeling, clinical application and conceptual representation. These groups were partially represented in both published frameworks. They were joined together to produce the combined set that was used in the current project. The first group of criteria is about the modeling in general and contains questions about type of evaluation (e.g. cost-effectiveness, retrospective) and the modeling method. This part of the framework was mainly taken from Cuzin et al. (1998). The BMJ framework requires only to identify the modeling approach. The second group is about clinical perspectives in general. The "study question" criteria was taken from the BMJ framework, and corresponds to "adopted perspectives" in Cuzin et al. (1998). The criteria to assess the meaning of "Results" is represented only in Cuzin et al. (1998) whereas the BMJ requires only to identify how study

results are represented. The last group of criteria is about the prostate cancer model. Alternatives and specification of utilities are discussed in both published frameworks. Cuzin et al.(1998) additionally required identification of the phases of the clinical process (e.g. screening and diagnosis, stage determination and post-treatment outcome) and population type (CUZIN et al., 1998).

Using this evaluation framework has highlighted the variability of published decision model representation. Omitting details on model structure, assumptions about data and structure in publication cause a "transparency" problem. This significantly reduces reusing models or parts of the models. To facilitate future reuse of the current model, a three level (structural, functional, and numeric) representation has been adopted as a guideline when considering model content.

The structural level shows how the model structure is coherent with the conceptual model of prostate cancer screening, diagnosis and management. Three separate conceptual models, screening and diagnosis, stage determination and post-treatment outcomes were initially created. An evidence integration approach (described later in this Chapter) was used to integrate these models into a final single model. The detailed representation of this integrated model should allow use of this model or its parts in future research and corresponds to the first research objective.

The functional level allows a decision maker to estimate all possible alternative strategies (described earlier in this Chapter). In the current study six alternative prostate cancer screening scenarios were evaluated with primary evidence coming from data of the Bayer clinical study. Alternative strategy represented general population undergoing tPSA and DRE based screening and primarily based on literature data was used in a sensitivity analysis to study the integrated model behavior. These scenarios correspond to the second

research objective. Even though many researchers have used heterogeneous sources of information for decision models, the formalism of this approach has not yet been described. The Cochrane collaboration organization has developed criteria for evidence selection and integration of different studies in order to increase significance of results. In a meta-analysis results from several studies are combined based on similarity of population, intervention studied, control and comparison groups , and patient outcome.

The link between meta-analysis and decision analysis was previously investigated by Lehman et al. (2000) and Petitti (1994) (LEHMANN et al., 2000, PETITTI, 1994). Our current study develops these observations and describes a formal method for evidence integration using a decision analysis model with structural, functional and numeric considerations enabling combination of parts from other models. Population, Intervention and Outcome characteristics were used to integrate different parts of the decision model. Decision model elements (chance nodes) were assigned to these three parameters. Integration is possible if there is a match between the decision model elements of the separate models (e.g. chance nodes have same attributes or represent the same population, intervention and outcome). This approach fills the gap between systematic review and meta-analysis from one side and decision analysis from another. The approach for evidence integration was successfully presented at the annual meeting of the American Association of Medical Informatics (AMIA) (GRANT and MOSHYK, 2002).

Contribution of previously published prostate cancer models

The first two models on prostate cancer screening were published in 1992. Mold et al. (1992) evaluated the benefit of using DRE based screening vs. no screening (MOLD et al., 1992). Launois (1992) evaluated six screening combinations of using DRE, TRUS and tPSA (LAUNOIS, 1992).

The question of whether tPSA based screening is beneficial or not (screening vs. no screening) was evaluated by some authors using a decision modeling approach in the middle of the 1990s. A single episode of screening was evaluated by Kranh et al. (1994), whereas Littrup et al. (1994b) and Cantor et al. (1995) evaluated a longitudinal screening algorithm (CANTOR et al., 1995, KRAHN et al., 1994, LITTRUP et al., 1994b).

These models have produced opposing results to date. The gain in life expectancy was shown for the screened patients, however when the strategies were compared by a quality of life scale, the decision was in favor of not screening. Critique revealed lack of standardized questionnaires, the variability in the gathering of preferences from different population groups (health practitioners, healthy volunteers, prostate cancer patients) (CANTOR et al., 1995, KRAHN et al., 1994, LAUNOIS, 1992, LITTRUP et al., 1994b, MOLD et al., 1992).

Two recent studies from the Quebec population have also given controversial results. Labrie et al. (1999) suggested that early detection and early curative treatment permit the decrease in mortality of prostate cancer patients (LABRIE et al., 1999). However Perron et al. (2002) suggested other reasons of recent decrease in mortality due to changes in disease management and the hormonal treatment of advanced disease (PERRON et al., 2002).

The first model that compared the new marker based strategies was published by Ellison et al. (2002). The tPSA, f/tPSA in conjunction with tPSA, and cPSA based strategies were evaluated. DRE was not taken into consideration (ELLISON et al., 2002). The author compared alternative strategies by the amount of unnecessary biopsies, costs and the chance of missing cancer cases.

The screening and diagnosis phase of the current model is a development of the Ellison et al.'s model (ELLISON et al., 2002). In comparison to his model, where the results

are based on a single episode of screening, the current work extends the evaluation of new marker strategies for repeated longitudinal prostate cancer screening. It is the first Markov model that evaluates use of new markers. The dataset used here was also used for Ellison's model combined with their own data.

In comparison to our study, authors of other models tend to use different cut offs for the markers and make a decision based on costs or other parameters first, without taking into account variation in diagnostic performance. Ellison et al. (2002) specifically studied several cut offs for cPSA based screening with different combinations of sensitivity/specificity in order to evaluate a complex utility function based on proportion of false positive and false negative results seeking to relate patient preference to diagnostic performance (ELLISON et al., 2002). In the current study it was decided to compare costs based on the same analytical performance of the different strategies. The discussion about using the ROC curve approach to adjust strategies to the same sensitivity level is presented later in this chapter.

The current model evaluated the same strategies as the Ellison et al. (2002) alternatives (ELLISON et al., 2002). In addition, we evaluated c/tPSA in conjunction with a tPSA+DRE based strategy. These methods (c/t and f/tPSA) have been suggested in the literature as competitive alternatives. The tPSA with/without DRE strategy, studied in the current model, appeared initially in the Littrup et al. (1994b) model. The latter study compared the impact of DRE on diagnostic performance during serial screening. It was found that performance of DRE decreases with every round of serial screening (LITTRUP et al., 1994b). Later Smith et al. (1996) also suggested that DRE may be avoided according to the results of a longitudinal study on prostate cancer markers (SMITH et al., 1996). In the current research cPSA and tPSA based strategies were studied with and without DRE as an additional test.

Utility assessment

Almost all authors have used the cost for screening program evaluation. Other parameters of choice have been either biopsy based parameters (efficacy) or quality of life (effectiveness).

Gustafsson et al. (1995a) used cost for evaluation of alternative screening strategies of a single screening episode. He calculated organizational direct and indirect costs and used these parameters to estimate ratios such as cost per cancer detected, and cost per detected treatable, local or small cancer (GUSTAFSSON et al., 1995a). Coley (1997a) provided evaluation of a single episode of screening vs. no-screening using a cost per saved life year parameter (COLEY, 1997a). Many researchers have used quality adjusted life years (QALY) parameters to estimate the effectiveness of the screening programs (e.g. Fleming et al. (1993), Littrup et al. (1994b), etc) but these parameters vary in the different studies. However no studies exist that compare the QALY evaluation between different modeling studies and hence the QALY parameter cannot be considered as interchangeable between the studies. Littrup et al. (1994b) estimated evolution of cost per cancer detected for repeated annual screening (FLEMING et al., 1993, LITTRUP et al., 1994b).

Two authors published models where costs were not a primary evaluation parameter. Draisma et al. (2003) evaluated mean lead time and rates of over-detection due to screening (DRAISMA et al., 2003). Grover et al. (2000a) evaluated clinical burden for prostate cancer with respect to disease specific mortality (GROVER et al., 2000a).

Commonly an evaluation of screening strategies using new markers is provided by comparison of how many biopsies can be avoided. This biopsy based utility might however lead to biased results. For example, number of biopsies per cancer detected can be represented as (FP+TP)/TP, where FT is a false positive result (biopsy is ordered but

appeared as negative), TP is a true positive result (identified cancer cases). This parameter is 1/PPV, where PPV is positive predictive value. Positive predictive value is a measure of diagnostic performance which depends on cancer prevalence. Since population variability is often observed between studies on prostate cancer screening (e.g. populations of healthy people, referred patients, patients from urology clinics, other facilities), it might be erroneous to use this parameter to make an inference from one study to another.

The cost, ratio of cost per cancer detected and overall mortality parameters were used in the current study. The expenses of a screening program were estimated using local hospital costs for a 5 year longitudinal program.

A clinical specialist (urologist) and laboratory staff were contacted about the price of diagnostic procedures. Total cost for each screening strategy was calculated per patient. In order to provide more robust results, Monte-Carlo simulation was used for cost utility estimation. The resulting utility was represented in the form of distributions. The utility function distributions were compared for all screening alternatives.

Results of model application

Two principle applications of the model were studied. The screening and diagnosis phase of the model was used to investigate the comparative costs of different screening strategies particularly those incorporating new markers. Secondly, the integrated model was studied to investigate the importance of key variables towards determining eventual patient outcome.

For reasons discussed above the markers for the different strategies were adjusted to comparable sensitivity. The costs of a repeated annual screening policy for each of these strategies was evaluated by Monte Carlo simulation based on the Markov model.

This simulation revealed no significant differences in costs between using either tPSA or cPSA biochemical markers. With the same diagnostic performance, it was found that DRE based strategies significantly increase the cost in comparison to the non-DRE based strategies. These findings support previous suggestions that DRE is economically expensive as a first line test for prostate cancer screening, assuming appropriate selection of biochemical marker cut-off (SMITH et al., 1996). Using adjusted sensitivities to ensure comparability of analytical performance, an economical benefit between new and classical biochemical marker strategies could not be demonstrated. This analysis supports the view that the use of biochemical markers as a screening test should be carefully controlled as to relative analytical performance.

The second application concerned sensitivity analysis using the integrated model. The key variables tested were the components of staging as occurs in the second phase of the model versus outcome in terms of expected mortality found in the third phase of the model. The importance of bone scan as part of staging and in relation to the biochemical markers remains controversial as earlier discussed. One particular alternative with bone scan and lymph nodectomy was included in the second phase model. This part of the model uses tPSA to separate the patients with higher and lower extent of advanced disease. The former group undergoes bone scan for exclusion of M1 cancer cases. Currently not included in the model are other imaging tools such as PET scan or computer tomography. Gleason score can also be used with or without tPSA in order to separate patients who undergo imaging, however only limited evidence is available for other imaging diagnostic tests and in using the Gleason score in the staging phase to satisfy modeling needs.

Also left out of the model at this stage although sufficient evidence is available for modeling concerns patients with organ confined disease who undergo pelvic lymph

nodectomy prior to prostatectomy which if positive for prostate cancer pelvic lymph nodes curtails the operation. Frazier (1994) has suggested that patients with limited node-positive (less than three nodes involved) disease selected for radical prostatectomy experience a survival advantage over those denied such therapy independently of adjunctive therapy (FRAZIER et al., 1994). According to Ouden and Schroeder (1998) survival is not significantly different between patients with locally confined prostate cancer and patients with locally advanced non-metastatic prostate cancer (T3) (OUDEN and SCHROEDER, 1998). It is possible that applying modeling to this and like evidence will assist in determining which is the better strategy.

Sensitivity analysis was used in the current study to test our second study hypothesis. The Tornado diagram (Figure 25) is a visual means of representation of the relative impact of the different staging variables probability estimates on the population overall mortality outcome. Post-treatment survival of T1/T2 groups is seen as the most sensitive parameter in a population of patients with newly diagnosed prostate cancer. This supports the hypothesis of Rientbergen et al. (1999) on how the specific stage mortality reduction can influence the patients' overall mortality (RIETBERGEN et al., 1999). Identification of data for survival of patients with T1/T2 with different treatment approaches including watchful waiting would allow to simulate the model around these datasets and test statistical significance of survival differences.

The second important parameter identified is the survival of M1 cancer patients. It is a relatively small group of people where the distant metastases can be identified at time of diagnosis, but the outcome of this sub-population is important for overall survival for prostate cancer patients.

Study limitations

Study limitations are considered with respect to three aspects (structural, functional representation, and data).

The model structure developed for the clinical stage determination and post-treatment phases was sufficient to answer the research objectives of this study. It does not at present contain cost evaluation data.

Limitations of functional representation relate to the number of alternatives considered. For the screening, phase uniform strategies for all 5 years of the hypothetical screening program have been considered. However change of strategy might occur for example DRE + biochemical marker at the 1st year, and a biochemical marker based program at follow-up visit. Screening strategy might be adjusted according to the available previous screening results. Accommodating these alternatives would require a large longitudinal dataset to create a decision model that provides precise estimation of patient population stratification based on previous screening results.

Commonly employed alternative has been reproduced for stage determination and post-treatment phases. However other alternatives, for example use of other possible imaging techniques, have not been considered in part because of the few data yet available.

Limitations of data are the following. Recent studies have reported that a population exposed to serial prostate cancer screening can change the distribution of prostate cancer presentation (RIETBERGEN et al., 1999). Such changes are not accommodated in the current study, where clinical characteristics (e.g. distribution of biochemical tests, DRE) were assumed to be constant during the first and follow-up screening. The possible changes might include decreasing DRE positive results and prevalence of prostate cancer characteristics.

Publication bias may also be present when combining results from similar studies during meta-analysis or aggregation. A funnel plot can reveal bias associated with reporting of positive effects of an intervention assuming that smaller studies with negative effects are under-reported. The horizontal axis of a funnel plot consist of measures of treatment effects (e.g. relative risk for treatment-effect studies, logarithm of the odds ratio). The vertical axis usually reports the sample size of the studies included in a meta-analysis. When all the studies have been located, the distribution of points should resemble a funnel. Other causes of bias other than publication bias might however give an equivalent plot [24].

Averaging of literature derived probabilistic information weighted by sample size was suggested by Lehmann et al. (2000) as one of the methods for evidence synthesis. Using Chi-square statistic confirmed a heterogeneity of pooled probabilities. The same authors have used another method, which is a Bayesian meta-analytic technique. Lehmann et al. (2000) used a hierarchical Bayesian model, similar to the random effect model in traditional meta-analysis. This method would be appropriate in our study because of heterogeneity of pooled probabilistic information. It incorporates prior belief and computes a wider confidence interval around estimates in the population parameters (LEHMANN et al., 2000). Another approach is to regroup selected publications in order to attain homogeneity of the population.

The biopsy was also assumed as a definitive answer as to whether the patient has prostate cancer or not. Several authors have however reported that 10-11% patients who had a negative biopsy result, were found to have prostate cancer on second biopsy soon after the first one (DJAVAN et al., 2001).

---

[24] http://www.cochrane.dk/cochrane/handbook/8_analysing_and_presenting_results/8.11_special_topics.htm Accessed Dec 5, 2004.

Modeling approaches

We found it useful to use both decision tree and influence diagram approaches, benefiting from this facility of the Reason Edge software. Two data sources were used to indirectly specify the second phase of modeling (preoperative stage determination). The tPSA distribution and clinical stage from the Bayer clinical study dataset were combined with probabilistic information about bone scan and lymph nodectomy results from published literature. The influence diagram representation allowed to calculate the distribution of patients between the different treatment groups. The Bayer clinical study data were also used to determine screening alternatives and calculate the distribution of patients between newly diagnosed cancer and no cancer groups after screening.

The software Data 3.5 provides an advantage of easy cost estimation, where a decision maker can attach costs (or utilities) to any branch (not only terminal nodes), and the application calculates the overall expected utility. This advantage was useful while performing sensitivity analysis on probabilistic parameters of the decision model. The Data 3.5 application also enables Monte-Carlo simulation.

According to expected utility theory, the winning scenario is chosen after comparing expected utilities. Applying the model once, expected utility is calculated as a single number, without assessment of the degree of confidence. When the Monte-Carlo approach is used, the expected utility is a distribution. Statistical methods can be used to compare distributions and provide confidence levels for similarity or difference. The notion of confidence level is more familiar to health care givers and corresponds to the EBM approach, where scientific evidence has to be supported by the quantitative measures.

Every round of prostate cancer screening was represented as a cycle in Markov model. After costs per each diagnostic test or procedure were attached to the decision tree,

the decision model provides cost of screening for 1 year. In order to calculate the cost per 5 year, the Markov model was executed for 5 cycles. This approach also accommodated well the post-treatment survival and longitudinal cost evaluation.

Other model approaches can be used. If the research question is about the evaluation of effectiveness parameter (e.g. quality adjusted life years) and the life expectancy is known for every post-treatment alternative, then a Markov model is not appropriate. QALY utility can be placed at the terminal nodes without consideration of the Markov states. Since the research hypothesis considered cost and efficacy evaluation over repeated screening, our model required use of the Markov modeling approach.

Sensitivity analysis is a further way to evaluate a robustness of results in addition to the Monte-Carlo simulation. During sensitivity analysis probabilistic and/or utility parameters are varied simultaneously or one at a time. The result of sensitivity analysis is often shown as a chart with axes representing dependency of expected utility from the varied parameter(s). As the decision maker systematically examines the influential parameters, the amount of charted data increases progressively. The Tornado diagram was used in the current study as a convenient way to represent the results of sensitivity analysis on several parameters. Parameters for sensitivity analysis were arbitrarily varied across a +/- 50% range.

The Tornado diagram is not the only reported method to estimate post-hoc robustness of decision problem to multiple parameters' estimates. Felli and Hazen (2004) presented an alternative approach to tornado diagrams. Javelin diagrams allow to conduct a probabilistic sensitivity analysis on multiple parameters simultaneously as well as does the tornado approach. However a decision maker may assign distributions to uncertain parameters and compute the probability of decision changes (FELLI and HAZEN, 2004).

Several authors have applied the ROC curve approach for the evaluation of diagnostic programs which include multiple tests (BAKER, 2000, ETZIONI et al., 2003, Mcintosh and PEPE, 2002, MURTAUGH, 1995). The objective for these studies has been to define appropriate combination of diagnostic tests (number of tests, cut offs) based on the ROC curve. The user can define desirable levels of sensitivity and specificity through assignment of weights and the decision is based on the optimal area under ROC curve. These authors have proposed analytical solutions to accommodate a different number of diagnostic tests (equations with multiple variables) and various outcome formats (e.g. nominal, discrete or real values). In comparison to such studies, this research project has proposed a less general and more feasible application of the ROC curve approach with less complex equations. First of all, a target overall sensitivity level was stated. Then cut offs of all tests except the last one were defined by the diagnostic protocol. The cut off for this last test was calculated as described earlier (see pp. 88-89). The calculations were then applied to 2 types of diagnostic protocols (DRE+BiochemicalMarker and DRE+BiochemicalMarker1+ BiochemicalMarker2).

Improvements to decision modeling

Specifying probabilistic information is the most time consuming part of decision modeling. This leads to the development of specialized models with specific assumptions, with limited chance of reuse and in consequence a lack of fully described models in printed scientific journals. More concerted attention to methods of evidence representation and evidence selection are two general ways to improve the decision modeling process.

Most medical journals now have an electronic version. Subscribers can access the text of the article or its abstract on-line. Examples of publication of anonymized clinical data on-line can be found at www.hutchon.freeserve.co.uk/demo.htm (HUTCHON, 2001). Some

journals require simultaneous submission of data for third party verification (ALTMAN and CATES, 2001). Raw clinical data can provide usually more information than summary results in final article. However there are several issues about confidentiality and property of the data that have to be solved.

Improvements in evidence selection from bibliographic databases would be facilitated by the possibility to make a search in structured abstracts taking advantage of the structure of the article such as population, intervention and outcome. An example is the work of Ida Sim on the structured presentation of articles of cancer trials (SIM et al., 2000).

A reasonable approach should be an integration of the decision model application and data sources. Assessment of the model in relation to data of clinical practice would also favor an automation of probabilistic data evaluation. For example during the process of model building a domain expert could select the population and thresholds from the database by defining queries to the database. Created once, the links between nodes and data source could be used many times as new information appears in the database. One could say that the preferable use of decision modeling software would be on top of a health care database (e.g. regional, hospital or large research study database).

Regarding the improvements in modeling results representation, Sanders et al. (2000) described a system for automatic clinical guidelines generation based on textual description associated with decision model elements. The author developed a web-based system, that creates an annotated flowchart algorithm automatically by analyzing an underlying decision model. This system constructs Clinical Practice Guidelines that represent the optimum policy as determined by the evidence and values reflected in the decision model. Users can view the evidence on which the guideline is based, review model inputs, or interrogate and re-run them using a web-based interface (SANDERS et al., 2000).

Decision model structure is usually defined by domain experts. An alternative way is to learn the structure of decision models from data. Structure learning from the database could be represented either as a decision tree or as a belief network. In general machine learning, algorithms have been used for feature selection in databases for some time. The current situation is that few software vendors enable interoperability between applications that could generate decision trees or belief networks and decision modeling applications. At this moment this has been done only for the integration of the data mining package Clementine and Hugin. Recently a group of researchers has shown the results of integration using WinMine (dependency networks generation) and MSBNx Toolkit (decision analysis application). But practically it has not been implemented yet. No one of the "off-the-shelf" software supports the whole process of model development (data source selection, data linking, export results findings).

As part of the early work during the current project a data mining approach was evaluated. Freely available for research purposes, the data mining software (J.Cheng's Belief Network PowerConstructor) provides an ability to graphically represent dependencies between clinical variables through learning Bayesian belief network structure from an investigational dataset. Data mining on clinical data set gives the graphical structure, which is similar to one based on the evidence from published clinical studies. The results of this experience were presented at the AMIA Annual meeting 2002 (GRANT and MOSHYK, 2002). The structure of the network elements represent the significant predictors of outcomes. A belief network was created from the clinical dataset. The outcome to be predicted was the biopsy result. It was found that the belief network from the predictive model and the decision model have graphical similarities but the nodes did not have the same

attributes of number of outcomes or conditional dependency links. Adapting automatically generated models requires domain experts' assistance.

Conclusion

Different methodologies (systematic review, clinical data analysis and decision modeling,) were examined for evaluation of prostate cancer screening using an evidence integration approach. The modeling process was based on creating a graph of a conceptual model of the problem, selecting probabilistic information based on the structural relationship between decision model elements and employing representation guidelines to avoid transparency problems and increase model reuse. Different questions were asked of the model. Examination of new marker screening strategies from a cost-detection perspective did not confirm an advantage of new over existing markers in a 5 year screening program. Sensitivity analysis of the integrated model suggested that optimised detection and therapy of early stage cancer should improve survival. This experience supports wider use of modeling for study of health care problems.

VI. Acknowledgments

I would like to acknowledge and thank the following people:

Dr. Andrew Grant,

Dr. François Bernard

Dr. Michel Carmel

M. Larbi Djaider,

Dr. Teophile Niyonsenga,

M. Philippe Caron,

M^{e}. Tu Hoang,

M^{e}. Carol Cheli,

and all members of the Clinical Science Program at Université de Sherbrooke for the support in this project.

VII. Appendicies

Table 29. Summary table on decision models on prostate cancer published between 1992-2003 (part 1)

Study	Type of evaluation	Objectives	Population	Methods	Alternatives
1. (CANTOR et al., 1995).	Retrospective Effectiveness	Evaluate efficacy of screening for prostate cancer	Men after 50 years old	Markov model Data from the literature and life-table data. Utility were determined by 10 married couples. Subjects evaluated outcomes of the treatment and treatment complications.	Screening vs. no screening
2. (COLEY, 1997a, COLEY, 1997b)	Retrospective cost/effectiveness study	Determine benefit for diff. population groups for one screening strategy (DRE+tPSA(4µg/l)	3 groups of patients (50-59, 60-69, 70-79)	Markov model; 6 months cycle, Data from published studies. Systematic review for 1966-1995. Non repeated screening.	Screening vs. diagnostics comparison between diff. populations for one screening strategy (DRE+tPSA(4µg/l)
3. (DRAISMA et al., 2003).	Retrospective Efficacy study	Determine the impact on the net benefits of screening	General population of men 55-74 y.o.	Monte-Carlo simulation of Markov process model; data from the European Randomized Study of Screening for Prostate Cancer	4 variations for a base population model
4. (ELLISON et al., 2002).	Retro/pro-specitive; cost-benefit analysis	Determine the appropriate tPSA derivate with biomarker's cut-off estimation for a screening	Referred patients with normal DRE under 75 y.o.	Decision tree, data from multi-institutional database; sensitivity analysis	5 alternatives (tPSA, f/tPSA, cPSA with various cut-offs)
5. (ETZIONI et al., 1996)	Retrospective, Efficacy study	Compare expected survival benefits using an age-specific tPSA cut offs	General population of men eligible for a prostate cancer	Markov model, probabilistic information from published studies	Two alternatives: tPSA based screening using 4 µg/l cut-off vs. using age-adjusted cut-offs

			screening		
6. (FLEMING et al., 1993).	Retrospective Cost/effective-ness analysis	Determine the impact of initial therapy on localized prostate cancer outcomes	Men of 60-75 years old	Markov model; probabilistic information from literature; sensitivity analysis	Therapy vs. watchful waiting
7. (GOTTLIEB et al., 1996).	Retrospective; Cost/effective-ness analysis	Determine utility for different excess tPSA levels as a function of age	Surgical candidates PSA tPSA 0-20 µg/l	Decision tree. The probability of a positive biopsy by evaluation of 71 consecutive patients at local institution. Disease-specific mortality rate from the literature	immediate prostate biopsy vs. no biopsy of the prostate at more than 70 years age
8. (GROSSFELD, 2000).	Retrospective Effectiveness study	Determine preferred management of a positive surgical margins after radical prostatectomy	Men undergone prostatectomy with positive surgical margins	Decision tree Literature and institution based estimates for probability of undetectable tPSA after treatment, complications. Utilities (from 0 to 1) were assigned by panel of experts Sensitivity analysis to determine threshold values	Immediate adjuvant radiation vs. surveillance with delayed radiation as necessary
9. (GROVER et al., 2000a).	Retrospective Demographic forecast	Forecast the health care requirements and outcomes associated with prostate cancer	General population	Markov model Probabilistic data from the literature Model computes life-expectancy, estimates the annual probability of a diagnosis of PC, progression to metastatic disease, death from PC and from other causes with or without previously diagnosed prostate cancer Validation with published observations for various cohorts of men	Undergoing screening population
10. (GROVER et al., 2000b).	Retrospective Economic forecast	Forecast the health care requirements and outcomes associated with prostate cancer	General population	Markov model Canadian age-specific incidence data; initial treatment choice for National Cancer Database The utilization rates of diagnostic tests and procedures for staging from US data. Model validated according to clinical output (see earlier)	Undergoing screening population

11. (GUSTAFSSON et al., 1995a).	Prospective, cost/efficacy study	Determine most cost/effective strategy	Men 55-70 y.o.	Decision tree Not repeated screening Randomly selected patients from clinical study	6 strategies of screening
12. (HILLNER et al., 1995).	Retrospective Cost/effective-ness analysis	Estimate cost-effectiveness of total androgen blockade with flutamide in M1 prostate cancer	M1 prostate cancer patients	Markov model Transitional probabilities from Intergroup 0036 study. Range for sensitivity analysis estimated form meta-analysis of similar studies Sensitivity analysis	Using flutamide vs. other studies results
13. (JAGER et al., 2000)	Retrospecitve Cost/ effectiveness study	Determine the appropriate use of MRI for preoperative staging of prostate cancer	Patients who were considered surgical candidates on the basis of clinical staging	Decision tree Data from published clinical studies Sensitivity analysis	Two strategies: radical prostatectomy with/without previous MRI examination
14. (KATTAN et al., 1997).	Retrospective Effectiveness study	Determine optimal management strategy for men who have localized prostate cancer	Men with clinically localized prostate cancer	Reuse Fleming decision model Two groups of patients (historical and prospective control) were used to elicit utilities. Then these utilities were applied to Fleming model. Monte-Carlo simulations to define group-level utilities.	Radical prostatectomy vs. watchful waiting
15. (KRAHN et al., 1994).	Retrospective; cost/effectiveness study	Estimate benefits and cost of different screening strategies and treatment of cancer with stage A-B	Stage A-B, men 50, 60, 70 y.o.	Markov model, Data from published clinical studies Efficacy data – most favourable data for screening Effectiveness data are based on False Positive and False Negative cases and complications and side effects from the treatment Sensitivity analysis Not repeated screening strategy	4 strategies of screening
16. (LAUNOIS, 1992).	Retrospective; cost/efficacy study	Determine more or less cost/effective strategies	Men >50 y.o.	Decision tree, Data from published clinical studies Not repeated PC screening Sensitivity analysis	6 strategies for screening

17. (LITTRUP et al., 1994a).	Retrospective Cost/benefit analysis	Evaluate economic performance of various tPSA screenings approaches	Population of men from ACS-NPCDP study	Exact formula ROC analysis to define appropriate biomarkers' cut-offs The cost value of benefits accrued due to earlier detection and treatment. Probability estimates, sensitivity, specificity of detection were based on completed data of clinical study ACS-NPCDP Sensitivity analysis	Screening vs. not screening
18. (LITTRUP et al., 1994b).	Prospective cost/efficacy study	Determine screening strategy for biopsy reduction with minimal loss in cancer detection	Men volunteers 55-70 y.o.	Decision tree Data from clinical study ACS NPCDP ROC analysis on tPSA cut offs Sensitivity analysis	9 strategies
19. (MOLD et al., 1992).	Retrospective, Effectiveness study	Evaluate value of performing or not performing periodic rectal examination	Asymptomatic 65 y.o. men	Decision tree Published data Utilities were determined by two primary care physicians using the Kaplan-Anderson Quality of Well-Being Scale	Screening with DRE vs. no screening
20. (MENG and CARROLL, 2000).	Retrospective Efficacy study	Determine thresholds for pelvic lymph node dissection before radical retropubic prostatectomy	Patients selected for prostatectomy	Decision tree, data from published studies. Utilities were determined by a panel of experts. Sensitivity analysis to determine important variables and calculate threshold values.	2 strategies (perform or omit lymph nodes)
21. (OGAWA and KATO, 1998)	Effectiveness study	Estimated the benefits of treatment strategies for early stage of prostate cancer	Patients with DRE negative prostate cancer under 70 years of age	Decision tree	Selective treatment vs. radical prostatectomy for all patients vs. watchful waiting for all patients
22. (ROSS et al., 2000).	Retrospective Cost/efficacy study	Compare prostate cancer mortality, tPSA testing rate, and biopsy rates	Men 50-70 years old	Markov model Probabilities from population data and surgical series. Monte-Carlo simulation of the natural history of prostate cancer grow. Sensitivity analysis	Various tPSA threshold for prostate biopsy, tPSA testing intervals, start age for tPSA testing

23. (SEIDENFELD et al., 1999)	Retrospective Cost/effectiveness study	Evaluate alternative strategies for androgen suppression as treatment of advanced prostate cancer	Patients with advanced prostate cancer stages	Markov model Probabilistic information from published studies Meta-analysis	Various combinations of androgen suppression treatment
24. (WOLF et al., 1993b).	Retrospective Efficacy study	Determine criteria for laparoscopic pelvic lymphadenectomy prior to radical prostatectomy	Prostate cancer patients undergoing pelvic lympadenectomy	Decision tree Utilities based on patient's preferences for different outcomes. Threshold analysis	Laparoscopic lymph node biopsy before operation or operation without pelvic lymph node biopsy
25. (WOLF et al., 1995).	Retrospective Cost/Efficacy analysis	Determine needs for imaging with computerized tomography or MRI and fine needle aspiration in the assessment of pelvic lymph nodes	Underwent the MRI (selected for prostatectomy) patients	Decision tree Diagnostic accuracy parameters from local hospital database. Data from published studies. Sensitivity analysis. Effectiveness parameters depends on the false/true positive/negative results of staging diagnostic tests	MRI imaging or no MRI imaging
26. (YOSHIMURA et al., 1998).	Retrospective; Effectiveness study	Evaluate the usefulness of pre-treatment prediction of clinically significant or insignificant tumor	T1c cancer patients detected with tPSA and negative DRE	Decision tree Utility depends on true/false positive/negative. Life expectancy and QALE Sensitivity analysis	3 strategies: prostatectomy; watchful waiting(WW); prostatectomy+WW

Table 30. Summary table on decision models on prostate cancer published between 1992-2003 (part 2)

Study	Phases	Utilities (primarily costs)	Results
1. (CANTOR et al., 1995).	Screening, treatment	No costs; QALY	Radiation was barely preferable to surgery of the early stages. The optimal treatment (if screening should occur) was relatively insensitive to variations in the model's values. The no-screening strategy was preferred to the screening strategy. The decision to screen is sensitive to changes in the patient's preferences regarding adverse effect of treatment The model evaluates the decision in the primary care settings rather than in a specialty clinic. Model did not incorporate short-term adverse effects on quality of life
2. (COLEY, 1997b).	Screening, treatment	Costs from Medicare database; cost/screened person; cost/saved life years	Result is sensitive to effectiveness criteria of treatment. Cost and marginal cost/effectiveness increases with age. Organisational costs not included.
3. (DRAISMA et al., 2003).	Screening, staging	No cost estimation; lead time and overdetection estimates	Mean lead times and rates of overdetection depends on a man's age at screening. Mean lead time and the overdetection rate were calculated for single screening, an annual or 4-year screening interval.
4. (ELLISON et al., 2002).	Screening	Office and staffing cost, assays and biopsy, indirect costs.	Authors have found dominant strategy (cPSA with cut-off of 3.8 μg/l) according to the cost-benefit ratio (cost vs. risk for false results detection). They also have used number of avoided biopsies for alternatives evaluation
5. (ETZIONI et al., 1996)	All phases	Expected survival benefits	Using a bound of 4.0 μg/l for all ages is more efficient than age-adjusted cut-offs. Average years of life saved per subject screened using tPSA > 4 μg/l were comparable to using the age-specific bound. Average years of life saved per cancer case were greater for tPSA > 4 μg/l than for age-specific.
6. (FLEMING et al., 1993).	Treatment	QALY, utility is based on the quality of life by treatment outcomes	The choice of watchful waiting is a reasonable alternative to invasive treatment for many men with localized prostatic carcinoma.
7. (GOTTLIEB et al., 1996).	Screening	Costs were based on charges at local hospital and were considered over a 2-year time frame. The financial costs considered were the direct costs of diagnostic testing and treatment. Marginal	Immediate prostate biopsy is not cost-effective in patients more than 70 years of age at all excess tPSA levels between 0-20 μg/l Costs were considered over only a 2 year time-frame rather than over the lifetime of the patient because of the considerable uncertainty in projecting out these costs over the life times of the patients.

		effectiveness in QALYs Marginal cost-effectiveness in $ per QALY of immediate prostate biopsy at diff. excess tPSA levels	
8. (GROSSFELD, 2000).	Treatment	No costs	Immediate radiation may be appropriate for patients with a positive surgical margins and a high likelihood of recurrent local rather than distant disease (low to intermediate grade, no evidence of seminal vesicle invasion and multiple positive margins)
9. (GROVER et al., 2000a).	All phases	no cost	The 10-years disease specific survival. Modelisation study. There are no comparison between alternatives. Survival forecast
10. (GROVER et al., 2000b).	All phases	The cost of initial and follow-up cancer therapies and complications included the costs of initial hospital services, physician fees and outpatient services (like disease monitoring and palliative care expenditures). Quebec and Ontario reimbursement schedules.	Direct medical cost due to prostate cancer for Canadian population of men between 40-80 years old. Modelisation study. Economic forecast
11. (GUSTAFSSON et al., 1995a).	Screening	Standard cost Cost of time spent by staff, material amortization cost, cost occupied rooms Organisational and indirect cost; cost/detected PC; cost/detected treatable PC; cost/detected local PC; cost/small cancer (<1.5cm)	TRUS+tPSA(4 µg/l) – 1st choice TRUS+DRE+tPSA(4 µg/l) – 2nd choice Complications due to biopsy T3 stage is considered as treatable Discussable way of the indirect costs valorization

12. (HILLNER et al., 1995).	Treatment	Cost were considered from a third-party-payer perspective (by Medicare or by VA)	Flutamide has an incremental cost-effectiveness more favorable than most accepted therapies. If drug costs are covered under health care reform, flutamide should be initiated and covered for all good performance status patients.
13. (JAGER et al., 2000)	Stage determina. tion	Cost of diagnostic procedures and treatment, QALY	A strategy to use MRI is cost-effective for men with moderate or high prior probability of extracapsular disease
14. (KATTAN et al., 1997).	Treatment	QALE, no cost	Some men will have better QALEs with radical prostatectomy and others with watchful waiting (a quality-based treatment benefit for radical prostatectomy for younger men and treatment harm for older men). The optimal treatment for the individual depends on his own unique preferences. A group-level recommendation cannot be substituted for an individual patient's decision regarding the management of localized prostate cancer
15. (KRAHN et al., 1994).	Screening, treatment	Standard hospital and ambulatory costs from Medicare database (adjusted cost); cost/year of saved life, marginal cost/saved life years. Cost of DRE = 0 (assumption)	Low gain in saved life for screening strategy Result is very sensitive to treatment efficacy Annual repetition cannot significantly change results No precise tPSA cut off
16. (LAUNOIS, 1992).	Screening	From database of Sécurité Sociale; cost/screened person, cost/detected prostate cancer case	3 strategies were found cost-effective out of 6 (those are able to detect 30%, 67% and 95% of cancers Results are sensitive to changes in diagnostic sensitivity/specificity of tests Cost/screened person is not an interesting criteria when efficacy is different. Lowering cost can advantage low effective costs. No precise cut off for tPSA
17. (LITTRUP et al., 1994a).	Screening; treatment	Estimated cost of all manipulations, complications Marginal cost, benefit-cost	If minimized future expenditures for terminal cancer care via decreases in therapy choices or coverage, no economic benefit for screening exists. The greatest immediate cost control issue is the marked increase in prostate cancer detection in the oldest age groups who have the least likelihood of mortality or morbidity benefits. Current cost savings may be possible with improved public health education about the appropriateness of early detection in the oldest age groups or those with significant pre-existing medical conditions.

18. (LITTRUP et al., 1994b).	Screening	Not clear valorisation of costs; cost/detected cancer Evolution of ratio in time(3 Successive annual screening)	tPSA with/without DRE – more cost/eff. strategy DRE+tPSA(4µg/l) – more effective Increasing cost/detected cancer with new successive screening Cost/benefit analysis No statistical analysis of diff. strategies Age distribution of volunteers is not explicit.
19. (MENG and CARROLL, 2000).	Staging	Efficacy (utility by expert opinion); indirect estimation of cost by efficacy parameters	Model favors omitting pelvic lymph node dissection. Outcomes would be equivalent for range of incidence of positive lymph nodes between 18% and 80%. Results are insensitive to the pelvic lymph node dissection complication rate
20. (MOLD et al., 1992).	Screening, staging, treatment	QALY, life expectancy	Periodic rectal examinations in asymptomatic men in the primary care setting does not lead to significant improvement in life expectancy and adversely affects quality of life. According to sensitivity analysis if more than 49% of prostate nodules are cancerous, then full evaluation and treatment is favored. Screening strategy does not correspond to real life
21. (OGAWA and KATO, 1998)	Treatment	Quality of life	A selective treatment with a reasonable specificity and sensitivity for detecting clinically significant cancer is more beneficial to the DRE negative prostate cancer patients under 70 - 80 years of age than radical prostatectomy or watchful waiting assigned to all patients. Quality-of-life is the most important factor for deciding the optimal treatments of prostate cancer
22. (ROSS et al., 2000).	Screening	No monetary cost. Authors used ration between number of tPSA or biopsy tests and cancer deaths prevented (proxy of monetary costs/prevented deaths); Effectiveness: number of PC deaths prevented. Incremental cost-effectiveness ration: number of additional tPSA test or biopsies to prevent 1 additional PC death	If screening for prostate cancer is beneficial, a screening strategy at age 40 and 45 years with a 2-year testing interval after age 50 years may be both more effective and require less testing than the standard strategy of annual tPSA testing beginning at age 50 years. Model did not take into account the morbidity from PC treatment and the effect of diagnosis and treatment on quality of life, the full cost associated with detection and management of prostate cancer

23. (SEIDENFELD et al., 1999)	Treatment	Cost-effectiveness analysis from societal perspective	For newly diagnosed prostate cancer patients with locally advanced or asymptomatic metastatic disease, the evidence is insufficient to determine whether primary androgen suppression initiated at diagnosis improves outcomes
24. (WOLF et al., 1995).	Staging	Marginal costs per patient benefited; efficacy as the appropriateness of a given outcome	Imaging would be beneficial if probability of nodal metastases were 45%. Cross-sectional pelvic imaging before radical prostatectomy is not justified routinely.
25. (WOLF et al., 1993b).	Staging	Efficacy, no costs	The treatment threshold prior to radical perineal prostatectomy is 20% (7-27% range). If the risk of pelvic lymph node metastases is 20% or greater, laparoscopic pelvic lymph nodectomy prior to intended radical perineal prostatectomy would be preferred.
26. (YOSHIMURA et al., 1998).	Staging; Treatment	QALE for outcomes; no costs	A selective treatment strategy of "Prostatectomy + Watchful Waiting" based on pre-treatment prediction for significant tumor is a beneficial alternative to radical prostatectomy unselectively assigned to all patients at the T1c stage, if a reasonable accuracy in prediction is attained.

VIII. Bibliography

1. Albertsen P C. The positive yield of imaging studies in the evaluation of men with newly diagnosed prostate cancer: A population based analysis. Journal of Urology 2000; (163): 1138-1143.

2. Alsikafi N F. 3 Year biochemical disease free survival following radical prostatectomy in men with poorly differentiated prostate cancer. Journal of Urology 1998; (161): Supplement.

3. Altman D G, Cates C. Authors should make their data available. BMJ 2001; (323): 1069-1070.

4. Amling C L, Blute M L, Lerner S E, Bergstralh E J, Bostwick D G, Zincke H. Influence of prostate-specific antigen testing on the spectrum of patients with prostate cancer undergoing radical prostatectomy at a large referral practice. Mayo Clin Proc 1998; (73): 401-406.

5. Andreassen S, Riekehr C, Kristensen B, Schonheyder HC, Leibovici L. Using probabilistic and decision-theoretic methods in treatment and prognosis modeling. Artificial Intelligence in Medicine 1999; (15): 121-134.

6. Ataus S, Citci A, Alici B, Onder A U, Sonmezoglu K, Erozenci A, Solok V. The value of serum prostate specific antigen and other parameters in detecting bone metastases in prostate cancer. International Urology & Nephrology 1999; (31): 481-489.

7. Augustin H, Hammerer P, Blonski J, Graefen M, Palisaar J, Daghofer F, Huland H, Erbersdobler A. Zonal location of prostate cancer: significance for disease-free survival after radical prostatectomy? Urology 2003; (62): 79-85.

8. Babaian R J, Dinney C P, Ramirez E I, Evans R B. Diagnostic testing for prostate cancer detection. Society of Surgical Oncology, 46th Annual Cancer Symposium in Conjunction with Society of Head and Neck Surgeons 1993; (March 18-21, 1993, Los Angeles, CA, p. 35, 1993. , 1993.).

9. Bader P I A. Is a limited lymph node dissection an adequate staging procedure for prostate cancer? Journal of Urology 2002; (168): 514-518.

10. Baker S G. Identifying combinations of cancer markers for further study as triggers of early intervention. Biometrics 2000; (56): 1082-1087.

11. Benoit R M, Naslund M J. The socioeconomic implications of prostate-specific antigen screening. Urologic Clinics of North America 1997; (24): 451.

12. Bielza C, Shenoy P. A comparison of graphical techniques for asymmetric decision problems. Working Paper No. 271. 1996. School of Business, University of Kansas, Lawrence, KS.

13. Black W C, Nease R F, Jr., Welch H G. Determining transition probabilities from mortality rates and autopsy findings. Med Decis Making 1997; (17): 87-93.

14. Bolla M, Gonzalez D, Warde P, Dubois J B, Mirimanoff R. Improved survival in patients with locally advanced prostate cancer treated with radiotherapy and goserelin. New England Journal of Medicine 1997; (337): 295-300.

15. Brawer M K. Screening for prostate cancer. Seminars in Surgical Oncology 2000; (18): 29-36.

16. Brawer M K, Cheli C D, Neaman I E, Goldblatt J, Smith C, Schwartz M K, Bruzek D J, Morris D L, Sokoll L J, Chan D W, Yeung K K, Partin A W, Allard W J. Complexed prostate specific antigen provides significant enhancement of specificity compared

with total prostate specific antigen for detecting prostate cancer. J Urol 2000; (163): 1476-1480.

17. Brawer M K, Ploch N R, Bigler S A. Prostate cancer tumor location as predicted by digital rectal examination transferred to ultrasound and ultrasound-guided prostate needle biopsy. Journal of Cellular Biochemistry - Supplement 1992; (16H): 74-77.

18. Brett T D. An analysis of digital rectal examination and serum-prostate-specific antigen in the early detection of prostate cancer in general practice. Family Practice 1998; (15): 529-533.

19. Briggs A, Sculpher M. An introduction to Markov modelling for economic evaluation. Pharmacoeconomics 1998; (13): 397-409.

20. Briggs A H. Handling uncertainty in cost-effectiveness models. Pharmacoeconomics 2000; (17): 479-500.

21. Buskirk S J, Pisansky T M, Atkinson E J, Schild S E, O'Brien P C, Wolfe J T, Zincke H. Lymph node-positive prostate cancer: evaluation of the results of the combination of androgen deprivation therapy and radiation therapy. Mayo Clinic Proceedings 2001; (76): 702-706.

22. Buxton M J, Drummond M F, van Hout B A, Prince R L, Sheldon T A, Szucs T, Vray M. Modelling in economic evaluation: an unavoidable fact of life. Health Econ 1997; (6): 217-227.

23. Candas B, Cusan L, Gomez J L, Diamond P, Suburu R E, Levesque J, Brousseau G, Belanger A, Labrie F. Evaluation of prostatic specific antigen and digital rectal examination as screening tests for prostate cancer. Prostate 2000; (45): 19-35.

24. Cantor S B, Spann S J, Volk R J, Cardenas M P, Warren M M. Prostate cancer screening - A decision analysis. Journal of Family Practice 1995; (41): 33-41.

25. Cao C G, Leong T Y. Learning conditional probabilities for dynamic influence structures in medical decision models. Journal of the American Medical Informatics Association 1997;848.

26. Cao C G, Leong T Y, Leong A P K, Seow F C. Dynamic decision analysis in medicine: a data-driven approach. International Journal of Medical Informatics 1998; (51): 13-28.

27. Catalona W J, Richie J P, deKernion J B, Ahmann F R, Ratliff T L, Dalkin B L, Kavoussi L R, MacFarlane M T, Southwick P C. Comparison of prostate specific antigen concentration versus prostate specific antigen density in the early detection of prostate cancer: receiver operating characteristic curves. [see comments]. Journal of Urology 1994; (152): 2031-2036.

28. Catalona W J, Smith D S, Wolfert R L, Wang T J, Rittenhouse H G, Ratliff T L, Nadler R B. Evaluation of percentage of free serum prostate-specific antigen to improve specificity of prostate cancer screening. [see comments]. JAMA 1995; (274): 1214-1220.

29. Cher D J, Lenert L A. Rapid approximation of confidence intervals for Markov process decision models: applications in decision support systems. J Am Med Inform Assoc 1997; (4): 301-312.

30. Chodak G W. Screening and early detection of prostate cancer. Cancer 1993; (71): 981-983.

31. Christensson A, Bjork T, Nilsson O, Dahlen U, Matikainen M T, Cockett A T, Abrahamsson P A, Lilja H. Serum prostate specific antigen complexed to alpha 1-antichymotrypsin as an indicator of prostate cancer. J Urol 1993; (150): 100-105.

32. Chybowski F M, Keller J J, Bergstralh E J, Oesterling J E. Predicting radionuclide bone scan findings in patients with newly diagnosed, untreated prostate cancer:

prostate specific antigen is superior to all other clinical parameters. J Urol 1991; (145): 313-318.

33. Claxton K, Neumann P J, Araki S, Weinstein M C. Bayesian value-of-information analysis. An application to a policy model of Alzheimer's disease. Int J Technol Assess Health Care 2001; (17): 38-55.

34. Coley C M M. Early detection of prostate cancer: Part I : Prior probability and effectiveness of tests. Annals of Internal Medicine 1997a; (126): 394-406.

35. Coley C M M. Early detection of prostate cancer: Part II : Estimating the risks, benefits, and costs. Annals of Internal Medicine 1997b; (126): 468-479.

36. Cooner W H, Mosley B R, Rutherford C L, Jr., Beard J H, Pond H S, Terry W J, Igel T C, Kidd D D. Prostate cancer detection in a clinical urological practice by ultrasonography, digital rectal examination and prostate specific antigen. Journal of Urology 1990; (143): 1146-1152.

37. Crawford E D, DeAntoni E P, Etzioni R, Schaefer V C, Olson R M, Ross C A. Serum prostate-specific antigen and digital rectal examination for early detection of prostate cancer in a national community-based program. The Prostate Cancer Education Council. Urology 1996; (47): 863-869.

38. Cuzin B, Maisonneuve H, Thoral F, Charvet-Protat S, Cordier H, Haslin N, Saul-Bertolone C. Opportunite d'un depistage systematique du cancer de la prostate par le dosage de l'antigene specifique de la prostate. 1998. ANAES/Service Évaluation Technologique.

39. DeMarzo A M, Nelson W G, Isaacs W B, Epstein J I. Pathological and molecular aspects of prostate cancer. Lancet 2003; (361): 955-964.

40. Dittmer S, Jensen F. Myopic value of information for influence diagrams. [Proc. of the 13th Conference on UAI], 142-149. 1997.

41. Djavan B, Ravery V, Zlotta A, Dobronski P, Dobrovits M, Fakhari M, Seitz C, Susani M, Borkowski A, Boccon-Gibod L, Schulman C. Prospective evaluation of prostate cancer detected on biopsies 1,2,3 and 4: When should we stop? Journal of Urology 2001; (166): 1679-1683.

42. Draisma G, Boer R, Otto S J, van der Cruijsen I, Damhuis R A, Schroeder F H, de Koning H J. 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): 868-878.

43. Drummond M F, Jefferson T O. Guidelines for authors and peer reviewers of economic submissions to the BMJ. The BMJ Economic Evaluation Working Party. BMJ 1996; (313): 275-283.

44. Eastham J A, Kattan M W. Disease recurrence in black and white men undergoing radical prostatectomy for clinical stage T1-T2 prostate cancer. [see comments]. Journal of Urology 2000; (163): 143-145.

45. Egawa S, Suyama K, Ohori M, Kawakami T, Kuwao S, Hirokado K, Hirano S, Yokoyama E, Uchida T, Koshiba K. Early detection of prostate cancer - Results of a prostate specific antigen-based detection program in Japan. Cancer 1995; (76): 463-472.

46. Ellison L, Cheli C D, Bright S, Veltri R W, Partin A W. Cost-benefit analysis of total, free/total, and complexed prostate-specific antigen for prostate cancer screening. Urology 2002; (60): 42-46.

47. Etzioni R, Kooperberg C, Pepe M, Smith R, Gann P H. Combining biomarkers to detect disease with application to prostate cancer. Biostatistics 2003; (4): 523-538.

48. Etzioni R, Shen Y, Petteway J C, Brawer M K. Age-specific prostate-specific antigen: a reassessment. Prostate Suppl 1996; (7:70-7.): 70-77.

49. Felli J C, Hazen G B. Javelin Diagrams: A Graphical Tool for Probablilistic Sensitivity Analysis. Decision analysis 2004; (1).

50. Fergany A. No difference in biochemical failures with or without pelvic lymph node dissection during radical prostatectomy: preliminary results. Journal of Urology 1999; (161): Supplement.

51. Fleming C, Wasson J H, Albertsen P C, Barry M J, Wennberg J E. A decision analysis of alternative treatment strategies for clinically localized prostate cancer. Prostate Patient Outcomes Research Team. JAMA 1993; (269): 2650-2658.

52. Fleshner N, Rakovitch E, Klotz L. Differences between urologists in the United States and Canada in the approach to prostate cancer. J Urol 2000; (163): 1461-1466.

53. Frazier H A, Robertson J E, Paulson D F. Does radical prostatectomy in the presence of positive pelvic lymph nodes enhance survival? World Journal of Urology 1994; (12): 308-312.

54. Freedland S J. Improved risk stratification for biochemical recurrence after radical prostatectomy using a novel risk group system based on prostate specific antigen density and biopsy Gleason score. Journal of Urology 2002; (168): 110-115.

55. Frydenberg M. Prostate cancer diagnosis and management. Lancet 1997; (349): 1681-1687.

56. Galen R S. Application of the predictive value model in the analysis of test effectiveness. Clin Lab Med 1982; (2): 685-699.

57. Garnick M B, Fair W R. Prostate cancer: emerging concepts. Part I. Ann Intern Med 1996a; (125): 118-125.

58. Garnick M B, Fair W R. Prostate cancer: emerging concepts. Part II. Ann Intern Med 1996b; (125): 205-212.

59. Gerber G S, Thisted R A, Scardino P T, Frohmuller H G, Schroeder F H, Paulson D F, Middleton A W, Jr., Rukstalis D B, Smith J A, Jr., Schellhammer P F, Ohori M, Chodak G W. Results of radical prostatectomy in men with clinically localized prostate cancer. JAMA 1996; (276): 615-619.

60. Glass T R, Tangen C M, Crawford E D, Thompson I. Metastatic carcinoma of the prostate: identifying prognostic groups using recursive partitioning. J Urol 2003; (169): 164-169.

61. Gleave M E, Coupland D, Drachenberg D, Cohen L, Kwong S, Goldenberg S L, Sullivan L D. Ability of serum prostate-specific antigen levels to predict normal bone scans in patients with newly diagnosed prostate cancer. Urology 1996; (47): 708-712.

62. Gottlieb R H, Mooney C, Mushlin A I, Rubens D J, Fultz P J. The prostate: decreasing cost-effectiveness of biopsy with advancing age. Invest Radiol 1996; (31): 84-90.

63. Gould D A, Birkmeyer J D. Efficacy versus effectiveness of carotid endarterectomy. Eff Clin Pract 1999; (2): 30-36.

64. Grant AM, Moshyk A. Integrating Evidence and Data into Decision Models for Ongoing Technology Evaluation. 2002. Washington, DC, American Medical Informatics Association. Proceedings of Annual meeting.

65. Grossfeld G D. Management of a positive surgical margin after radical prostatectomy: Decision analysis. Journal of Urology 2000; (164): 93-100.

66. Grover S A, Coupal L, Zowall H, Rajan R, Trachtenberg J, Elhilali M, Chetner M, Goldenberg L. The clinical burden of prostate cancer in Canada: forecasts from the Montreal Prostate Cancer Model. CMAJ 2000a; (162): 977-983.

67. Grover S A, Coupal L, Zowall H, Rajan R, Trachtenberg J, Elhilali M, Chetner M, Goldenberg L. The economic burden of prostate cancer in Canada: forecasts from the Montreal Prostate Cancer Model. CMAJ 2000b; (162): 987-992.

68. Gustafsson O, Carlsson P, Norming U, Nyman C R, Svensson H. Cost-effectiveness analysis in early detection of prostate cancer: An evaluation of six screening strategies in a randomly selected population of 2,400 men. Prostate 1995a; (26): 299-309.

69. Gustafsson O, Mansour E, Norming U, Carlsson A, Tornblom M, Nyman C R. Prostate-specific antigen (PSA), PSA density and age-adjusted PSA reference values in screening for prostate cancer--a study of a randomly selected population of 2,400 men. Scandinavian Journal of Urology & Nephrology 1998; (32): 373-377.

70. Gustafsson O, Theorell T, Norming U, Perski A, Ohstrom M, Nyman C R. Psychological reactions in men screened for prostate cancer. British Journal of Urology 1995b; (75): 631-636.

71. Guyatt G H, Tugwell P X, Feeny D H, Haynes R B, Drummond M. A framework for clinical evaluation of diagnostic technologies. CMAJ 1986; (134): 587-594.

72. Han M, Partin A W, Piantadosi S, Epstein J I, Walsh P C. Era specific biochemical recurrence-free survival following radical prostatectomy for clinically localized prostate cancer. Journal of Urology 2001; (166): 416-419.

73. Han M, Walsh P C, Partin A W, Rodriguez R. Ability of the 1992 and 1997 American Joint Committee on Cancer staging systems for prostate cancer to predict progression-free survival after radical prostatectomy for stage T2 disease. J Urol 2000; (164): 89-92.

74. Helfand M, Pauker S G. Influence diagrams: a new dimension for decision models. Med Decis Making 1997; (17): 351-352.

75. Higashihara E, Nutahara K, Kojima M, Okegawa T, Miura I, Miyata A, Kato M, Sugisaki H, Tomaru T. Significance of free prostate-specific antigen and gamma-seminoprotein in the screening of prostate cancer. Prostate Suppl 1996; (7): 40-47.

76. Hillner B E, Mcleod D G, Crawford E D, Bennett C L. Estimating the cost effectiveness of total androgen blockade with flutamide in M1 prostate cancer. Urology 1995; (45): 633-640.

77. Hoedemaeker R F, Rietbergen J B, Kranse R, Schroeder F H, van der Kwast T H. Histopathological prostate cancer characteristics at radical prostatectomy after population based screening. Journal of Urology 2000; (164): 411-415.

78. Huncharek M, Muscat J. Serum prostate-specific antigen as a predictor of radiographic staging studies in newly diagnosed prostate cancer. Cancer Investigation 1995; (13): 31-35.

79. Hutchon D J. Publishing raw data and real time statistical analysis on e-journals. BMJ 2001; (322): 530.

80. Jager G J, Severens J L, Thornbury J R, de la Rosette J J M C, Ruijs S H J, Barentsz J O. Prostate cancer staging: Should MR imaging be used? - A decision analytic approach. Radiology 2000; (215): 445-451.

81. Jager P L. Treatment with radioactive 89strontium for patients with bone metastases from prostate cancer. BJU International 2000; (86): 929-934.

82. Kattan M W, Cowen M E, Miles B J. A decision analysis for treatment of clinically localized prostate cancer. J Gen Intern Med 1997; (12): 299-305.

83. Keefer D L, Kirkwood C W, Corner J L. Perspective on decision analysis applications, 1990-2001. Decision analysis 2004; (1): 5-24.

84. Kemp P M. Which patients with prostatic carcinoma require a staging bone scan? British Journal of Urology 1995; (79): 611-614.

85. Kessler B, Albertsen P. The natural history of prostate cancer. Urol Clin North Am 2003; (30): 219-226.

86. Kessler J M. Decision analysis in the formulary process. Am J Health Syst Pharm 1997; (54 Suppl 1): S5-S8.

87. Kikuchi E, Nakashima J, Ishibashi M, Ohigashi T, Asakura H, Tachibana M, Murai M. Prostate specific antigen adjusted for transition zone volume: the most powerful method for detecting prostate carcinoma. Cancer 2000; (89): 842-849.

88. Kindrick A V, Grossfeld G D, Stier D M, Flanders S C, Henning J M, Carroll P R. Use of imaging tests for staging newly diagnosed prostate cancer: trends from the CaPSURE database. Journal of Urology 1998; (160): 2102-2106.

89. Kirby R S, Kirby M G, Feneley M R, McNicholas T, McLean A, Webb J A. Screening for carcinoma of the prostate: a GP based study. Br J Urol 1994; (74): 64-71.

90. Krahn M D, Mahoney J E, Eckman M H, Trachtenberg J, Pauker S G, Detsky A S. Screening for prostate cancer. A decision analytic view. JAMA 1994; (272): 773-780.

91. Krahn M D, Naglie G, Naimark D, Redelmeier D A, Detsky A S. Primer on medical decision analysis: Part 4--Analyzing the model and interpreting the results. Med Decis Making 1997; (17): 142-151.

92. Kristensen F, Horder M, Poulsen P. Health Technology Assessment Handbook. 2001. Danish Institute for Health Technology Assessment.

93. Kwok Y S, Kim C, Heidenreich P A. Medical therapy or coronary artery bypass graft surgery for chronic stable angina: an update using decision analysis. Am J Med 2001; (111): 89-95.

94. Labrie F, Candas B, Cusan L, Gomez J L, Diamond P, Suburu R, Lemay M. Diagnosis of advanced or noncurable prostate cancer can be practically eliminated by prostate-specific antigen. Urology 1996; (47): 212-217.

95. Labrie F, Candas B, Dupont A, Cusan L, Gomez J L, Suburu R E, Diamond P, Levesque J, Belanger A. Screening decreases prostate cancer death: first analysis of the 1988 Quebec prospective randomized controlled trial. Prostate 1999; (38): 83-91.

96. Lau A H, Leong T Y. PROBES: A framework for probability elicitation from experts. Journal of the American Medical Informatics Association 1999;301-305.

97. Launois R. Cost-effectiveness analysis of strategies for screening prostatic cancer. Dev Health Econ Public Policy 1992; (1): 81-108.

98. Lee C T, Oesterling J E. Using prostate-specific antigen to eliminate the staging radionuclide bone scan. Urol Clin North Am 1997; (24): 389-394.

99. Lehmann H P, Hinton R, Morello P, Santoli J. Developmental dysplasia of the hip practice guideline: technical report. Committee on Quality Improvement, and Subcommittee on Developmental Dysplasia of the Hip. Pediatrics 2000; (105): E57.

100. Lein M, Jung K, Hammerer P, Graefen M, Semjonow A, Stieber P, Ossendorf M, Luboldt H J, Brux B, Stephan C, Schnorr D, Loening S A. A multicenter clinical trial on the use of alpha1-antichymotrypsin- prostate-specific antigen in prostate cancer diagnosis. Prostate 2001; (47): 77-84.

101. Lentini M, Marzano D, Perrone M, Annunziata S, Cianetti A. Comparison of free/total PSA (F/T PSA) ratio and PSA density (PSAD) in the early diagnosis of cancer of the prostate. Archivio Italiano di Urologia, Andrologia 1997; (69 Suppl 1): 101-104.

102. Leong T Y. Multiple perspective dynamic decision making. Artificial Intelligence 1998; (105): 209-261.

103. Leong T Y, Cao C. Modelling medical decisions in DynaMoL: a new general framework of dynamic decision analysis. Medinfo 1998; (9 Pt 1): 483-487.

104. Letaief B, Boughattas S, Hassine H, Essabbah H. Bone scan and prostate cancer. Ann Urol (Paris) 2000; (34): 254-265.

105. Levran Z, Gonzalez J A, Diokno A C, Jafri S Z, Steinert B W. Are pelvic computed tomography, bone scan and pelvic lymphadenectomy necessary in the staging of prostatic cancer? British Journal of Urology 1995; (75): 778-781.

106. Lin K, Szabo Z, Chin B B, Civelek A C. The value of a baseline bone scan in patients with newly diagnosed prostate cancer. Clinical Nuclear Medicine 1999; (24): 579-582.

107. Littrup P J, Goodman A C, Mettlin C J, Murphy G P. Cost analyses of prostate cancer screening: frameworks for discussion. Investigators of the American Cancer Society-National Prostate Cancer Detection Project. Journal of Urology 1994a; (152): 1873-1877.

108. Littrup P J, Kane R A, Mettlin C J, Murphy G P, Lee F, Toi A, Badalament R, Babaian R J. Cost-effective prostate cancer detection - Reduction of low-yield biopsies. Cancer 1994b; (74): 3146-3158.

109. Lorente J A, Valenzuela H, Morote J, Gelabert A. Serum bone alkaline phosphatase levels enhance the clinical utility of prostate specific antigen in the staging of newly diagnosed prostate cancer patients. European Journal of Nuclear Medicine 1999; (26): 625-632.

110. Luboldt H J, Husing J, Rubben H, Altwein J E, Bichler K H, Czaja D, Fornara P, Jockel K H, Schalkhauser K, Weissbach L, Wirth M. Early detection of prostate cancer in German urological practice by digital rectal examination and prostate-specific antigen. Urologe 2000; (39): 330-333.

111. Maeda H, Arai Y, Ishitoya S, Okubo K, Aoki Y, Okada T. Prostate specific antigen adjusted for the transition zone volume as an indicator of prostate cancer. Journal of Urology 1997; (158): 2193-2196.

112. Manyak M J, Javitt M C. The role of computerized tomography, magnetic resonance imaging, bone scan, and monoclonal antibody nuclear scan for prognosis prediction in prostate cancer. Seminars in Urologic Oncology 1998; (16): 145-152.

113. Matzkevich I, Abramson B. Decision-Analytic Networks in Artificial-Intelligence. Management Science 1995; (41): 1-22.

114. McIntosh M W, Pepe M S. Combining several screening tests: optimality of the risk score. Biometrics 2002; (58): 657-664.

115. McIntosh M W, Urban N, Karlan B. Generating longitudinal screening algorithms using novel biomarkers for disease. Cancer Epidemiol Biomarkers Prev 2002; (11): 159-166.

116. Mehta V, Kushniruk A, Gauthier S, Richard Y, Deland E, Veilleux M, Grant A. Use of evidence in the process of practice change in a clinical team: a study-forming part of the Autocontrol Project. Int J Med Inf 1998; (51): 169-180.

117. Meng M V, Carroll P R. When is pelvic lymph node dissection necessary before radical prostatectomy? A decision analysis. Journal of Urology 2000; (164): 1235-1240.

118. Mettlin C. The status of prostate cancer early detection. Cancer 1993; (72): 1050-1055.

119. Mettlin C, Lee F, Drago J, Murphy G P. The American Cancer Society National Prostate Cancer Detection Project. Findings on the detection of early prostate cancer in 2425 men. Cancer 1991; (67): 2949-2958.

120. Metz C E. Basic principles of ROC analysis. Semin Nucl Med 1978; (8): 283-298.

121. Miller M C, O'Dowd G J, Partin A W, Veltri R W. Contemporary use of complexed PSA and calculated percent free PSA for early detection of prostate cancer: impact of changing disease demographics. Urology 2001; (57): 1105-1111.

122. Mitchell I D, Croal B L, Dickie A, Cohen N P, Ross I. A prospective study to evaluate the role of complexed prostate specific antigen and free/total prostate specific antigen ratio for the diagnosis of prostate cancer. J Urol 2001; (165): 1549-1553.

123. Mold J W, Holtgrave D R, Bisonni R S, Marley D S, Wright R A, Spann S J. The evaluation and treatment of men with asymptomatic prostate nodules in primary care: a decision analysis. Journal of Family Practice 1992; (34): 561-568.

124. Morote J, Raventos C X, Lorente J A, Lopez-Pacios M A, Encabo G, de T, I, Andreu J. Comparison of percent free prostate specific antigen and prostate specific antigen density as methods to enhance prostate specific antigen specificity in early prostate cancer detection in men with normal rectal examination and prostate specific antigen between 4.1 and 10 ng./ml. Journal of Urology 1997; (158): 502-504.

125. Murtaugh P A. ROC curves with multiple marker measurements. Biometrics 1995; (51): 1514-1522.

126. Neheman A, Shotland Y, Metz Y, Stein A. Screening for early detection of prostate cancer (first experience in Israel). Harefuah 2001; (140): 4-10, 88, 87.

127. Nielsen T, Jensen V. Welldefined decision scenarios. 502-511. 1999. S.F., Cal., Morgan Kaufmann Publishers. 15th Conference on Uncertainty in Artificial Intelligence (UAI-99). 7-30-1999.

128. Nielsen T, Jensen V. Representing and Solving Asymmetric Bayesian Decision Problems. 416-425. 2000. Morgan Kaufmann Publishers. Proceedings of the Sixteenth Conference on Uncertainty in Artificial Intelligence.

129. O'Dowd G J, Veltri R W, Orozco R, Miller M C, Oesterling J E. Update on the appropriate staging evaluation for newly diagnosed prostate cancer. J Urol 1997; (158): 687-698.

130. Oesterling J E. Free, complexed and total serum prostate specific antigen: The establishment of appropriate reference ranges for their concentrations and ratios. Journal of Urology 1995; (154): 1090-1095.

131. Ogawa O, Kato T. Decision making analysis in the treatment of prostate cancer. Nippon Rinsho 1998; (56): 2087-2091.

132. Okegawa T, Noda H, Nutahara K, Higashihara E. Comparisons of the various combinations of free, complexed, and total prostate-specific antigen for the detection of prostate cancer. European Urology 2000; (38): 380-387.

133. Ouden D, Schroeder F H. The treatment of locally advanced (T3) prostatic carcinoma using radical prostatectomy or radiotherapy. A review. Tijdschrift voor Gerontologie en Geriatrie 1998; (29): 74-79.

134. Ozdal O L, Aprikian A G, Begin L R, Behlouli H, Tanguay S. Comparative evaluation of various prostate specific antigen ratios for the early detection of prostate cancer. BJU Int 2004; (93): 970-974.

135. Partin A W, Kattan M W, Subong E N, Walsh P C, Wojno K J, Oesterling J E, Scardino P T, Pearson J D. Combination of prostate-specific antigen, clinical stage, and Gleason score to predict pathological stage of localized prostate cancer. A multi-institutional update. JAMA 1997; (277): 1445-1451.

136. Partin A W, Yoo J, Carter H B, Pearson J D, Chan D W, Epstein J I, Walsh P C. The use of prostate specific antigen, clinical stage and Gleason score to predict pathological stage in men with localized prostate cancer. Journal of Urology 1993; (150): 110-114.

137. Pauker S G, Kassirer J P. The threshold approach to clinical decision making. N Engl J Med 1980; (302): 1109-1117.

138. Perron L, Moore L, Bairati I, Bernard P M, Meyer F. PSA screening and prostate cancer mortality. CMAJ 2002; (166): 586-591.

139. Petitti D B. Meta-analysis, decision analysis and cost-effectiveness analysis : methods for quantitative synthesis in medicine. New York : Oxford University Press, 1994.

140. Pizzoccaro M, Valtorta A, Cappellano F, Sironi D, Catanzaro F. Current role of echography, of PSA and of PSAD in the diagnosis of prostatic carcinoma. Archivio Italiano di Urologia, Andrologia 1994; (66): 77-80.

141. Porter C R, O'Donnell C, Crawford E D, Gamito E J, Sentizimary B, De Rosalia A, Tewari A. Predicting the outcome of prostate biopsy in a racially diverse population: a prospective study. Urology 2002; (60): 831-835.

142. Potter S R, Horniger W, Tinzl M, Bartsch G, Partin A W. Age, prostate-specific antigen, and digital rectal examination as determinants of the probability of having prostate cancer. Urology 2001; (57): 1100-1104.

143. Presti J C, Jr. Prostate cancer: assessment of risk using digital rectal examination, tumor grade, prostate-specific antigen, and systematic biopsy. Radiologic Clinics of North America 2000; (38): 49-58.

144. Presti J C, Jr., Chang J J, Bhargava V, Shinohara K. The optimal systematic prostate biopsy scheme should include 8 rather than 6 biopsies: results of a prospective clinical trial. Journal of Urology 2000; (163): 163-166.

145. Prestigiacomo A F, Stamey T A. Can free and total prostate specific antigen and prostatic volume distinguish between men with negative and positive systematic ultrasound guided prostate biopsies? J Urol 1997; (157): 189-194.

146. Qi R, Poole D. New method for influence diagram evaluation. Computational Intelligence 1995; (11): 498-528.

147. Recker F, Lummen G. Prostatic carcinoma. Screening--when and what? Therapeutische Umschau 2000; (57): 33-37.

148. Redelmeier D A, Detsky A S, Krahn M D, Naimark D, Naglie G. Guidelines for verbal presentations of medical decision analyses. Med Decis Making 1997; (17): 228-230.

149. Rietbergen J B, Hoedemaeker R F, Kruger A E, Kirkels W J, Schroeder F H. The changing pattern of prostate cancer at the time of diagnosis: characteristics of screen detected prostate cancer in a population based screening study. Journal of Urology 1999; (161): 1192-1198.

150. Rietbergen J B, Kruger A E, Hoedemaeker R F, BANGMA C H, Kirkels W J, Schroeder F H. Repeat screening for prostate cancer after 1-year followup in 984 biopsied men: clinical and pathological features of detected cancer. Journal of Urology 1998; (160): 2121-2125.

151. Roberts R O, Bergstralh E J, Peterson N R, Bostwick D G, Lieber M M, Jacobsen S J. Positive and negative biopsies in the pre-prostate specific antigen and prostate specific antigen eras, 1980 to 1997. J Urol 2000; (163): 1471-1475.

152. Ross K S, Carter H B, Pearson J D, Guess H A. Comparative efficiency of prostate-specific antigen screening strategies for prostate cancer detection. JAMA 2000; (284): 1399-1405.

153. Rubin M A, Fang M, Stifelman M D, Devries G M, Buttyan R, Benson M C, Katz A E, Olsson C A, Otoole K. Enhanced reverse transcriptase polymerase chain reaction for prostate-specific antigen combined with needle biopsy results - A superior predictor of PT3 disease. Molecular Diagnosis 1997; (2): 135-145.

154. Rudoni M, Antonini G, Favro M, Baroli A, Brambilla M, Cardani G, Ciardi L, Sacchetti G M, Inglese E. The clinical value of prostate-specific antigen and bone scintigraphy in the staging of patients with newly diagnosed, pathologically proven prostate cancer. European Journal of Nuclear Medicine 1995; (22): 207-211.

155. Rydh A, Tomic R, Tavelin B, Hietala S O, Damber J E. Predictive value of prostate-specific antigen, tumour stage and tumour grade for the outcome of bone scintigraphy in patients with newly diagnosed prostate cancer. Scandinavian Journal of Urology & Nephrology 1999; (33): 89-93.

156. Sanchez L A, Lee J T. Applied pharmacoeconomics: modeling data from internal and external sources. Am J Health Syst Pharm 2000; (57): 146-155.

157. Sanders G D, Nease R F, Jr., Owens D K. Design and pilot evaluation of a system to develop computer-based site-specific practice guidelines from decision models. Med Decis Making 2000; (20): 145-159.

158. Sargent D, Allegra C. Issues in clinical trial design for tumor marker studies. Semin Oncol 2002; (29): 222-230.

159. Sarma A V, Schottenfeld D. Prostate cancer incidence, mortality, and survival trends in the United States: 1981-2001. Semin Urol Oncol 2002; (20): 3-9.

160. Scardino P T, Shinohara K, Wheeler T M, Carter S S. Staging of prostate cancer. Value of ultrasonography. Urol Clin North Am 1989; (16): 713-734.

161. Seidenfeld J, Samson D J, Aronson N, Albertson P C, Bayoumi A M, Bennett C L, Brown A, Garber A, Gere M, Hasselblad V, Wilt T, Ziegler K. Relative effectiveness and cost-effectiveness of methods of androgen suppression in the treatment of advanced prostate cancer. Evid Rep Technol Assess (Summ) 1999;i-246, I1.

162. Shapiro A, Lebensart P D, Pode D, Bloom R A. The clinical utility of transrectal ultrasound and digital rectal examination in the diagnosis of prostate cancer. British Journal of Radiology 1994; (67): 668-671.

163. Sim I, Owens D K, Lavori P W, Rennels G D. Electronic trial banks: a complementary method for reporting randomized trials. Med Decis Making 2000; (20): 440-450.

164. Simpson K N. Problems and perspectives on the use of decision-analysis models for prostate cancer. J Urol 1994; (152): 1888-1893.

165. Smith D S, Catalona W J, Herschman J D. Longitudinal screening for prostate cancer with prostate-specific antigen. JAMA 1996; (276): 1309-1315.

166. Smyth P. Belief networks, hidden Markov models and Markov random fields: A unifying view. Pattern Recognition Letters 1997; (18): 1261-1268.

167. Sohayda C, Kupelian P A, Levin H S, Klein E A. Extent of extracapsular extension in localized prostate cancer. Urology 2000; (55): 382-386.

168. Sonnenberg F A, Beck J R. Markov models in medical decision making: a practical guide. Medical Decision Making 1993; (13): 322-338.

169. Stamey T A, Sozen T S, Yemoto C M, McNeal J E. Classification of localized untreated prostate cancer based on 791 men treated only with radical prostatectomy: common ground for therapeutic trials and TNM subgroups. J Urol 1998; (159): 2009-2012.

170. Stokkel M P, Zwinderman A H, Zwartendijk J, Pauwels E K, Eck-Smit B L. The value of pretreatment clinical and biochemical parameters in staging and prognostic stratification of patients with newly diagnosed prostate carcinoma. International Journal of Biological Markers 1998; (13): 70-76.

171. Tanguay S, Begin L R, Elhilali M, Behlouli H, Karakiewicz P I, Aprikian A G. Comparative evaluation of total PSA, free/total PSA, and complexed PSA in prostate cancer detection. Urology 2002; (59): 261-265.

172. Thompson I M, Zeidman E J. Current urological practice: routine urological examination and early detection of carcinoma of the prostate. Journal of Urology 1992; (148): 326-329.

173. Toubert M E, Schlageter M H, Bron J, Teillac P, Le Duc A, Najean Y. Screening for cancer of the prostate using prostate-specific antigen. Presse Med 1990; (19): 1139-1142.

174. Varenhorst E, Carlsson P, Pedersen K. Clinical and economic considerations in the treatment of prostate cancer. Pharmacoeconomics 1994; (6): 127-141.

175. Vashi A R, Oesterling J E. Percent free prostate-specific antigen: entering a new era in the detection of prostate cancer. Mayo Clin Proc 1997; (72): 337-344.

176. Vijayakumar V, Vijayakumar S, Quadri S F, Blend M J. Can prostate-specific antigen levels predict bone scan evidence of metastases in newly diagnosed prostate cancer? Am J Clin Oncol 1994; (17): 432-436.

177. Vis A N. Does PSA screening reduce prostate cancer mortality? CMAJ 2002; (166): 600-601.

178. Waidelich R, Jansen H M, Stieber P, Schmeller N, Lamerz R, Werdan K, Fateh-Moghadam A. Screening for prostatic carcinoma with prostate specific antigen. Anticancer Research 1997; (17): 2979-2981.

179. Wang C G, Leong T Y. Knowledge-based formulation of dynamic decision models. Pricai'98: Topics in Artificial Intelligence 1998; (1531): 506-517.

180. Wingo P A, Tong T, Bolden S. Cancer statistics, 1995. CA Cancer J Clin 1995; (45): 8-30.

181. Wolf J S, Cher M, Dall'era M, Presti J C, Jr., Hricak H, Carroll P R. The use and accuracy of cross-sectional imaging and fine needle aspiration cytology for detection of pelvic lymph node metastases before radical prostatectomy. J Urol 1995; (153): 993-999.

182. Wolf J S, Shinohara K, Carroll P R, Narayan P. Combined role of transrectal ultrasonography, Gleason score, and prostate-specific antigen in predicting organ-confined prostate cancer. Urology 1993a; (42): 131-137.

183. Wolf J S, Shinohara K, Kerlikowske K M, Narayan P, Stoller M L, Carroll P R. Selection of patients for laparoscopic pelvic lymphadenectomy prior to radical prostatectomy: a decision analysis. Urology 1993b; (42): 680-688.

184. Woodrum D L. Interpretation of free prostate specific antigen clinical research studies for the detection of prostate cancer. Journal of Urology 1998; (159): 5-12.

185. Wymenga L F, Boomsma J H, Groenier K, Piers D A, Mensink H J. Routine bone scans in patients with prostate cancer related to serum prostate-specific antigen and alkaline phosphatase. BJU International 2001; (88): 226-230.

186. Yoshimura N, Takami N, Ogawa O, Kakehi Y, Okada Y, Fukui T, Yoshida O. Decision analysis for treatment of early stage prostate cancer. Japanese Journal of Cancer Research 1998; (89): 681-689.

187. Yu K K, Hricak H. Imaging prostate cancer. Radiologic Clinics of North America 2000; (38): 59-85.