Markers of Differentiation and Prognosis in Prostate Cancer
- a Morphological and Immunohistochemical Study

Axel Glaessgen
MARKERS OF DIFFERENTIATION AND PROGNOSIS IN PROSTATE CANCER
- A MORPHOLOGICAL AND IMMUNOHISTOCHEMICAL STUDY

Axel Glaessgen

Stockholm 2008
Doctoral Thesis
Markers of differentiation and prognosis in prostate cancer – a morphological and immunohistochemical study.

All previously published papers were reproduced with permission from the publisher.

Published and printed by Laserics Digital Print AB
Box 20082, SE-161 02 Bromma, Sweden
© Axel Glaessgen, 2008
ABSTRACT

Prostate cancer (PCa) is the most common malignancy in men worldwide. The disease shows a wide range of patient outcome between indolent and highly aggressive tumor behavior. Prognostic factors are needed to stratify PCa patients in different prognostic groups for treatment decision and to predict outcome. Today serum prostate specific antigen (PSA), TNM staging and histomorphological prognostic factors such as the Gleason score (GS) are used in clinical practice for prediction of prognosis.

Despite its powerful prognostic power, the utility of GS has decreased in the last decades due to a grade and stage shift. Efforts have been made to develop the GS and modified systems such as percent Gleason grade 4/5 (%GG4/5) and modified Gleason score (mGS) have been suggested. Almost monthly new potential biomarkers for PCa are reported but most of them will have only very limited clinical impact. To be useful, a marker must provide prognostic information, which is independent from that of established prognostic factors, its reproducibility must be satisfactory and the information obtained clinically relevant.

The interobserver reproducibility of the conventional GS, %GG4/5 and mGS was analyzed among four uropathologists. The overall reproducibility of %GG4/5 and mGS was at least as good as that of the GS. However, clustering of mGS in odd scores was found and severe disagreement was more common than with GS. The ability of prostate needle biopsies to correctly predict %GG4/5 in prostatectomy specimens was found to be almost as good as for GS.

We investigated the expression of pancreatic duodenal homeobox-1 (PDX-1) and heat shock proteins (HSP) 27, 60 and 70 in prostate tissue. Tissue microarrays (TMA) were constructed for analysis of prognostic value (289 PCas, median follow-up 48.9 months), expression in benign tissues, high-grade PIN, primary PCa and lymph node metastases. Two independent observers evaluated intensity and extent of immunohistochemical staining. HSP 27 and 60, but not PDX-1 and HSP 70, correlated with biochemical recurrence. In a multivariate analysis including histological prognostic factors HSP 60 was an independent predictor of recurrence. PDX-1 was overexpressed in cancer vs. benign tissue, but also in atrophy and high-grade PIN. PDX-1 decreased with higher Gleason pattern and in metastases. There was only slight interobserver agreement for extent of immunoreactivity of PDX-1 and HSPs but moderate to substantial agreement for intensity of the staining. Thus, the question was raised whether staining extent should be estimated on TMA. Presence of PDX-1 protein in benign and malignant prostatic tissue was confirmed by Western blot. In conclusion we suggest that HSP 27 and HSP 60 are predictors of biochemical recurrence after radical prostatectomy and PDX-1 is of potential interest in the pathogenesis of PCa.

To my parents Helga and Jochen Glaessgen,
my wife Daria Glaessgen
and our children Ioana and Felix
LIST OF PUBLICATIONS


Interobserver reproducibility of percent Gleason grade 4/5 in total prostatectomy specimens.


II. Glaessgen A, Hamberg H, Pihl CG, Sundelin B, Nilsson B, Egevad L.

Interobserver reproducibility of percent Gleason grade 4/5 in prostate biopsies.

Journal of Urology 2004; 171(2 Pt 1): 664-7

III. Glaessgen A, Hamberg H, Pihl CG, Sundelin B, Nilsson B, Egevad L.

Interobserver reproducibility of modified Gleason score in radical prostatectomy specimens.

Virchows Archives 2004; 445(1): 17-21

IV. Glaessgen A, Busch C, Norberg M, Häggman M, Nilsson B, Egevad L.

Prediction of percent Gleason grade 4/5 by multiple core biopsies.


Expression of PDX-1 in prostate cancer, prostatic intraepithelial neoplasia and benign prostatic tissue.

Acta Pathologica, Microbiologica et Immunologica Scandinavica, in press


Heat shock proteins 27, 60 and 70 as prognostic markers of prostate cancer.

Acta Pathologica, Microbiologica et Immunologica Scandinavica, submitted
# TABLE OF CONTENTS

**List of abbreviations**

1 Background .................................................................................................................. 1
   1.1 The prostate .............................................................................................................. 1
      1.1.1 Anatomy ............................................................................................................ 1
      1.1.2 Histology .......................................................................................................... 2
      1.1.3 Physiology ......................................................................................................... 2
   1.2 Prostate cancer ........................................................................................................ 3
      1.2.1 Epidemiology .................................................................................................... 3
      1.2.2 Etiology and risk factors .................................................................................... 4
      1.2.3 Molecular genetics and cytogenetics ................................................................. 6
      1.2.4 Pathology of prostate cancer ............................................................................. 8
      1.2.5 Prognostic and predictive factors in prostate cancer ....................................... 12

2 Aims of the study ......................................................................................................... 31

3 Material and methods .................................................................................................. 32
   3.1 Reproducibility and prediction studies (papers I to IV) ......................................... 32
   3.2 Protein expression studies (papers V and VI) ....................................................... 33
      3.2.1 Material ............................................................................................................ 33
      3.2.2 Tissue microarray (TMA) construction ............................................................ 34
      3.2.3 Immunohistochemistry ..................................................................................... 35
      3.2.4 TMA immunostaining evaluation .................................................................... 35
      3.2.5 Western blot ..................................................................................................... 36
   3.3 Statistical analysis (papers I to VI) ........................................................................ 36

4 Results and discussion ................................................................................................. 38
   4.1 Reproducibility and prediction studies (papers I to IV) ......................................... 38
      4.1.1 Paper I and III: Interobserver reproducibility of %GG4/5 and mGS in prostatectomy specimens ................................................................. 38
      4.1.2 Paper II: Interobserver reproducibility of %GG4/5 in prostate biopsies ........... 39
      4.1.3 Paper IV: Prediction of %GG4/5 ...................................................................... 40
      4.1.4 Discussion papers I to IV ................................................................................. 40
   4.2 Protein expression studies (papers V and VI) ....................................................... 47
      4.2.1 Paper V: Pancreatic duodenal homeobox-1 ...................................................... 47
      4.2.2 Paper VI: Heat shock proteins 27, 60 and 70 .................................................... 49
      4.2.3 Reproducibility of TMA staining evaluation .................................................... 52

5 Acknowledgements ....................................................................................................... 55

6 References .................................................................................................................... 56
<table>
<thead>
<tr>
<th>Abbreviation</th>
<th>Definition</th>
</tr>
</thead>
<tbody>
<tr>
<td>%GG4/5</td>
<td>Percent Gleason grade 4/5</td>
</tr>
<tr>
<td>AR</td>
<td>Androgen receptor</td>
</tr>
<tr>
<td>DHT</td>
<td>Dihydrotestosterone</td>
</tr>
<tr>
<td>EPE</td>
<td>Extraprostatic extension</td>
</tr>
<tr>
<td>GS</td>
<td>Gleason score</td>
</tr>
<tr>
<td>HGPIN</td>
<td>High grade PIN</td>
</tr>
<tr>
<td>HSP</td>
<td>Heat shock protein</td>
</tr>
<tr>
<td>IOR</td>
<td>Interobserver reproducibility</td>
</tr>
<tr>
<td>k</td>
<td>Kappa statistic</td>
</tr>
<tr>
<td>LGPIN</td>
<td>Low grade PIN</td>
</tr>
<tr>
<td>mGS</td>
<td>Modified Gleason score</td>
</tr>
<tr>
<td>NMD</td>
<td>Nuclear morphometric descriptors</td>
</tr>
<tr>
<td>PCa</td>
<td>Prostate cancer</td>
</tr>
<tr>
<td>PDX-1</td>
<td>Pancreatic duodenal homeobox-1</td>
</tr>
<tr>
<td>PIA</td>
<td>Proliferative inflammatory atrophy</td>
</tr>
<tr>
<td>PIN</td>
<td>Prostatic intraepithelial neoplasia</td>
</tr>
<tr>
<td>PSA</td>
<td>Prostate specific antigen</td>
</tr>
<tr>
<td>PZ</td>
<td>Peripheral zone of the prostate</td>
</tr>
<tr>
<td>QNG</td>
<td>Quantitative nuclear grade</td>
</tr>
<tr>
<td>SVI</td>
<td>Seminal vesicle invasion</td>
</tr>
<tr>
<td>TMA</td>
<td>Tissue microarray</td>
</tr>
<tr>
<td>TURP</td>
<td>Transurethral resection of the prostate</td>
</tr>
<tr>
<td>TZ</td>
<td>Transition zone of the prostate</td>
</tr>
<tr>
<td>wk</td>
<td>Weighted kappa statistic</td>
</tr>
</tbody>
</table>
1 BACKGROUND

1.1 THE PROSTATE

1.1.1 Anatomy

The prostate is a sex related gland that belongs to the male reproductive system. Having the size of a walnut and a weight around 20 g, it is situated around the upper part of the urethra and lies with its superior surface (the base) immediately below the urinary bladder and with its inferior surface (the apex) above the pelvic musculofascial floor. A thin layer of connective tissue surrounds the prostate. The Denonvillier’s fascia separates its posterior surface from the rectum. The prostate is anteriorly fixed to the...
pubic bone with the puboprostatic ligaments, being held in the dorsal vein plexus between these structures. In addition there are two inferior-lateral surfaces (McNeal 1972; McNeal 1988; Nickel 1999).

Described first by McNeal, the zonal anatomy of the prostate includes the peripheral, transition and central zone, representing approximately 65%, 10% and 25%, respectively of the normal organ volume. The zonal anatomy is important in prostate pathology. Thus, most cancers develop in the Peripheral Zone (PZ) and benign hyperplasia mainly in the Transition Zone (TZ) of the prostate gland (McNeal 1988).

1.1.2 Histology

The prostate is composed of fibromuscular stroma and 30-50 glands that empty into the prostatic urethra. Two cell layers form the branching duct-acinar system: the luminal columnar secretory cells and the basal cells (McNeal 1972; Blacklock 1974; McNeal 1988; Nickel 1999). The architectural structure and polarity of the glandular cells are important for diagnosing and grading PCa.

The prostate does not have a well-defined capsule (Ayala, Ro et al. 1989), but the most peripheral layer of the fibromuscular stroma, bordering on periprostatic fat, forms a pseudocapsule. The organ border becomes very intricate at the base of the prostate, particularly at the junction with the seminal vesicles, while in other areas prostatic glands extend to the edge of the prostate without a distinct capsule, leading to difficulties in the assessment of extraprostatic extension (EPE).

1.1.3 Physiology

The main function of the prostate is to secrete a milky, alkaline fluid into the urethra during ejaculation. The prostate fluid is a semen component and represents one half to two thirds of ejaculate volume. It helps to nourish and protect the sperm during intercourse. It is slightly acid (pH 6.5) and contains several secretory products like acid phosphatase, citrate, zinc, soluble fraction proteins, carbohydrates, electrolytes, polyamines, hormones, lipids and growth factors (Fair and Cordonnier 1978; Weidner, Jantos et al. 1991; Zaichick, Sviridova et al. 1996). Among them is prostate-specific antigen (PSA), which proved to be of paramount importance for diagnosing PCa. PSA is a protease and its main role is to keep the semen liquid (Neal, Clejan et al. 1992;
Nickel 1999). Its role as a prognostic factor in PCa will be described in chapter 1.2.5.3.1.

The size and function of the prostate are regulated by hormones belonging either to the hypothalamic-pituitary-testicular or adrenal gland axis. Androgens with their major circulating form testosterone are of great importance for the growth and maintenance of the prostate gland. Free testosterone passes through the cell membrane into the prostatic epithelial cells where it is metabolized to dihydrotestosterone (DHT) by the enzyme 5-alpha-reductase. Within the glandular cells DHT binds to androgen receptors (AR) with a ten-fold greater affinity than testosterone. Subsequently, the AR-DHT complex targets within specific DNA sequences, known as androgen response elements and leads to activation of cell functions, including growth and proliferation (Carlson and Katzenellenbogen 1990; Deslypere, Young et al. 1992; Russell and Wilson 1994).

1.2 PROSTATE CANCER

1.2.1 Epidemiology

PCa is the most common malignancy in men worldwide. After lung cancer, PCa is the second leading cause of death in men in the industrialized countries. The precise worldwide prevalence of PCa is difficult to estimate because of lack of complete statistical data but according to the World Health Organization there were 679,023 new cases and 221,002 deaths from PCa worldwide in 2002 (Ferlay 2004). In 2004, PCa was estimated to account for 198,100 new cancer diagnoses and 31,500 deaths in the USA, about one death every 15 min (Cooperberg, Moul et al. 2005). Sweden is one of the countries with the highest PCa incidences in the world. PCa is the most common cancer in Swedish men with an annual incidence of approximately 9000 cases in 2003 and a mortality of 2352 men in 2002 (Cancerfonden 2003). In the Stockholm-Gotland region the number of newly diagnosed cases increased from 250 in the beginning of the 1960s to more than 2000 cases per year in 2005, while there was a continuous improvement of relative survival during this time period (relative five-year survival 2004 more than 80%) (Onkologiskt Centrum 2007).

In industrialized nations, PCa incidence is rising, while mortality is decreasing (Baade, Coory et al. 2004; Parkin, Bray et al. 2005). Decreasing PCa mortality is also seen in
the European community, with significant reductions in Germany, France, Spain, Italy, and the United Kingdom (Baade, Coory et al. 2004).

The increasing PCa incidence is partly due to successful early detection by the use of the serum PSA test, hereby narrowing the still enormous gap between clinical incidence (8-15% lifetime risk) and autopsy-based prevalence (80% by age 80 years) (Sakr, Haas et al. 1993).

Interestingly, there is a huge geographical variation of incidence (100-fold) and mortality (20-fold). The highest incidence is encountered in North America and Northwest Europe while the disease is rather rare in Asia, Africa and South America (Parker, Tong et al. 1997). The population with the highest incidence and mortality in the world are African American men in the United States while Asians show the lowest rates. Nevertheless these rates have rapidly raised in many Asian countries since the past two decades. This is supposed to be a result of not only enhanced detection but also westernization of the lifestyle, with increased obesity and consumption of fat (Pu, Chiang et al. 2004; Freedland and Isaacs 2005). The importance of dietary, socioeconomic and environmental factors is illustrated by the increasing risk of PCa in Asian immigrants in the USA (Whittemore, Kolonel et al. 1995). Genetical alterations probably also contribute to the huge geographical variation of PCa incidence. In conclusion, differences in risk by race may be due to one or more factors: genetics, exposure to carcinogens and life-style factors such as diet and decision-making (e.g. detection of cancer).

1.2.2 Etiology and risk factors

Despite all research efforts the causes of PCa remain unclear. It is believed that PCa has a multi-factorial origin with environmental as well as genetic factors, reflecting a complex pathogenesis.

Aging is the most significant risk factor (Abate-Shen and Shen 2000). The mean age at diagnosis is between 72 and 74 years and 85% of the patients are diagnosed after the age of 65. Before the age of 50 PCa is still rarely diagnosed though an increased incidence in younger men has been observed lately (Merrill, Potosky et al. 1996; Merrill and Lyon 2000; Gronberg 2003). An autopsy study showed that 8% of men in their twenties had invasive PCa while the prevalence increased to more than 75% after the age of 85 years. It was suggested that most men would get PCa if they lived long enough (Sakr, Haas et al. 1993; Sakr, Grignon et al. 1996)
A so-called “Western” diet, high in refined carbohydrates, saturated fats and calories and low in fruit and vegetables is now common not only in industrialized countries but also in many other countries (Flegal, Carroll et al. 2002; Peytremann-Bridevaux, Faeh et al. 2007). However, data showing an association between Western diet and PCa are inconsistent (Hsing, Deng et al. 2000; Wu, Hu et al. 2006), reflecting the complex multifactorial nature of this issue. Men in developing countries who adopted a Western diet had an increased PCa risk (Hsing, Deng et al. 2000). More recently, the specific type of dietary fat was shown to be important, with saturated fat increasing the risk (Whittemore, Kolonel et al. 1995; Kristal, Cohen et al. 2002) whereas polyunsaturated fats might have a protective effect (Bidoli, Talamini et al. 2005).

High levels of lycopene and carotenoids which proved to have anti-oxidative capacity are contained in high levels in tomatoes and their high intake may correlate with a risk reduction for PCa development (Basu and Imrhan 2007). However, a European prospective study was not able to show a relationship between plasma carotenoids, retinol and tocopherols and the overall risk of PCa (Key, Appleby et al. 2007). Other protective dietary components such as vitamin B12, folate, vitamin E and D, selenium and zinc were discussed but further research is demanded (Johansson, Appleby et al. 2008).

“Western” diet contributed to marked increases in the percentage of overweight and obese men in many industrialized nations over the last few decades. Studies investigating the PCa risk of obese men produced conflicting results. While some found an association between increased Body-Mass-Index (BMI) and increased risk for PCa (Andersson, Wolk et al. 1997; Veierod, Laake et al. 1997; Putnam, Cerhan et al. 2000; Engeland, Tretli et al. 2003) others found no association (Nilsen and Vatten 1999; Schuurman, Goldbohm et al. 2000). A recent meta-analysis of prospective cohort studies concluded that obesity was associated with a significant, but weak, increased PCa risk (MacInnis and English 2006). Prospective cohort studies consistently found an association between increased BMI and risk of PCa mortality, suggesting that obesity might play a more important role in PCa progression than in its initiation (Andersson, Wolk et al. 1997; Rodriguez, Patel et al. 2001; Calle, Rodriguez et al. 2003). Obesity could be linked to PCa by means of other mechanisms than diet, for example altered hormonal levels, increased serum levels of estradiol, insulin, free IGF-1, and leptin. The relevance of these changes is that all of these steroid and peptide hormones were linked with PCa (Giovannucci 2003; Baillargeon and Rose 2006).
Another disease associated to altered levels of insulin, testosterone, leptin and IGF-1 is diabetes mellitus. A meta-analysis in diabetic men including 19 studies performed by Kasper et al. showed a 16% decrease of the risk of developing PCa (Kasper and Giovannucci 2006). There are several possible explanations of this. Insulin, which is positively associated with proliferation of both normal and cancerous prostate cells showed decreased levels in diabetic men. Elevated leptin levels were positively associated to risk for advanced PCa while decreased levels were found in prolonged hypoinsulinemic states. Hypoinsulinemia caused upregulation of IGF-binding protein 1, thus limiting the bioavailability of circulating IGF-1, which at high levels is a risk factor for PCa. Low testosterone levels in diabetic men may be protective (Shaneyfelt, Husein et al. 2000).

Inflammation is another factor that is possibly involved in PCa development (De Marzo, Platz et al. 2007). Many others risk factors for PCa have been studied, like other dietary components, the influence of occupation, sexual and physical activity and social factors but more studies are needed to evaluate their possible role (Bostwick, Burke et al. 2004).

1.2.3 Molecular genetics and cytogenetics

A variety of molecular and cytogenetic abnormalities are associated with PCa. Although certain loci are related to familial PCa incidence, disease heterogeneity impeded the discovery of predisposition genes (Amanatullah, Reutens et al. 2000). The most frequent sites of genetic material loss in PCa are on chromosomes 13q, 8p, 6q, 5q, 16q, 18q, 2q, 4q, 10q, and Y (in decreasing order). The most frequent gains are seen on chromosomes 8q, 17q, Xq, 7q, 3q, 9q, 1q, and Xp (Alers, Rochat et al. 2000). Gain of genetic material on chromosome 8 is the most frequent numeric anomaly in PIN and PCa (Emmert-Buck, Vocke et al. 1995), suggesting that alterations of this chromosome and/or a tumor suppressor gene(s) on the short arm may be important for the initiation or early progression of PCa. Chromosomal loss on 8p22 with concurrent gain on 8q associates with adverse disease outcome (Macoska, Trybus et al. 2000). Loss at 8p21, the site of the prostate-specific homeobox gene NKX3.1, also correlates with tumor progression (Bowen, Bubendorf et al. 2000). Hypermethylation of the GSTP1 promotor is the most commonly described epigenetic alteration in PCa. GSTP1 is involved in the detoxification of many xenobiotics as a part of the cellular protection system against toxic effects (De Marzo, Meeker et al. 2003).
Recently a new gene fusion in PCa was discovered: the fusion between the androgen regulated gene TMPRSS2 (21q22.3) and one of the ETS genes: ERG (21q22.2), ETV1 (7p21.2) or ETV4 (17q21) (Tomlins, Rhodes et al. 2005). Among these, the TMPRSS2-ERG fusion is the most common (Tomlins, Mehra et al. 2006), occurring in up to 50% of clinically localized PCas in hospital-based cohorts (Perner, Demichelis et al. 2006). The high incidence of PCa probably makes this fusion the most common genomic alteration in human cancers so far described. Common somatic genetic and epigenetic changes in PCa are listed in Table 1.
1.2.4 Pathology of prostate cancer

1.2.4.1 Progression from precancerous lesions to prostate cancer

The pathogenesis of PCa is still not completely elucidated. It is considered that transition from benign prostate over preneoplastic lesions to PCa is a continuous process (Bostwick and Brawer 1987). Several potential precursor lesions of PCa are described such as low-grade prostatic epithelial neoplasia, atrophy and atypical adenomatous hyperplasia but no data are as convincing as those published for high-grade prostatic epithelial neoplasia (HGPIN) (Montironi, Mazzucchelli et al. 2007).

1.2.4.1.1 Proliferative inflammatory atrophy (PIA)

Inflammation is a common finding in prostate specimens. Many lesions containing inflammation are associated with atrophic epithelium or focal epithelial atrophy. The term proliferative inflammatory atrophy (PIA) was introduced by De Marzo (De Marzo, Marchi et al. 1999) to express the increased fraction of epithelial cells that appear to be proliferating in focal atrophy. PIA is associated with acute and chronic inflammation caused by a variety of potential agents (Figure 3). Proliferation in the setting of longstanding chronic inflammation seems to predispose to carcinoma in different organs and may even play a possible role in the pathogenesis of PCa. PIA may emerge as a consequence of epithelial damage caused by oxidative stress most likely derived from surrounding inflammatory cells. Many of the molecular and genetic changes seen in HGPIN and PCa are also documented in PIA. Morphological studies observed merging of focal atrophy with HGPIN and a close relation to early carcinoma (De Marzo, DeWeese et al. 2004). PIA, HGPIN and PCa are multifocal and share...
similar locations in the prostate zones (McNeal 1988). Many of the proliferating cells in PIA have an immature secretory cell phenotype similar to that in PIN and PCa (De Marzo, Marchi et al. 1999). GSTP1 is responsible for detoxifying carcinogens and

Figure 3: Possible causes of prostate inflammation. (Reprinted by permission from Macmillan Publishers Ltd: Nature Reviews Cancer 7(4): 256-69, copyright 2007)

Inflammatory oxidants. Hypermethylated (i.e. inactivated) GSTP1 was reported in around 6% of PIA lesions, 70% of PIN lesions and in >90% of PCas, while no hypermethylation was observed in benign and hyperplastic prostate tissue (Lee, Morton et al. 1994; Nakayama, Bennett et al. 2003). Immunophenotypic features of PIA are increased expression of GSTP1 and Bcl-2 and decreased p27KIP1 expression.

Further investigations are needed to determine whether inflammation, atrophy and oxidative stress are responsible for prostate carcinogenesis (De Marzo, Platz et al. 2007).
1.2.4.1.2 Prostatic intraepithelial neoplasia (PIN)

In the early 1960s McNeal postulated that PCa arises from active ductal/acinar epithelium and introduced the term “intraductal dysplasia of the prostate” (McNeal 1965). This concept was not widely accepted until 1986 when McNeal and Bostwick described morphological criteria and established a 3-tier grading system that predicted the association with invasive carcinoma (McNeal and Bostwick 1986). Later on the term prostatic intraepithelial neoplasia (PIN) was introduced (Bostwick and Brawer 1987) and the 3-tier grading system was merged into two grades to avoid reproducibility problems. Low-grade (LG) PIN included grade 1 and high-grade (HG) PIN included grade 2 and 3. PIN is defined as an intra-acinar epithelial proliferation with significant nuclear atypia in the secretory cells (Bostwick 1988). HGPIN is classified in four architectural patterns: tufting, micropapillary, cribriform and flat.

HGPIN features most of the phenotypic and biochemical changes in cancer without invasion of the basal membrane. The linkage between PCa and HGPIN is based on clinicomorphological, epidemiological, genetic and molecular evidence. With advancing age the incidence and extent of HGPIN increases. HGPIN correlates with the presence, frequency, severity and extent of PCa (Sakr, Grignon et al. 1996; Bostwick 2000). HGPIN and PCa are both multifocal and share similar locations in the prostate zones (Qian, Wollan et al. 1997; Montironi, Mazzucchelli et al. 2000). Transition from HGPIN to PCa can be observed morphologically and they share several cytological features (Bostwick 1989). Rates of cell proliferation and apoptosis are increased in both HGPIN and PCa (Colombel, Symmans et al. 1993; Montironi, Galluzzi et al. 1993). Phenotypically and morphometrically HGPIN and PCa are similar (Vis and Van Der Kwast 2001). Some genetic and molecular alterations are common in both lesions (Qian, Bostwick et al. 1995). The basal cell layer is disrupted in HGPIN and absent in PCa (Lipski, Garcia et al. 1996). In both entities neovascularization is increased compared to normal prostate tissue (Montironi, Diamanti et al. 1996).

The recognition of HGPIN is clinically important because of its strong association to PCa and the finding of isolated HGPIN in prostate biopsies should prompt the clinician to perform repeated biopsies (Gokden, Roehl et al. 2005).
1.2.4.2 Morphology

The vast majority of PCa’s are common acinar adenocarcinomas accounting for more than 90% of the cases (Cookson 2001). Morphological adenocarcinoma variants like ductal, mucinous, pseudohyperplastic, foamy gland and small cell cancers contribute to less than 5% of PCas. The prognosis of several of these variants is unclear because they are rare and usually combined with conventional PCa but some of them have a particular aggressive clinical behavior and should be distinguished from conventional prostatic adenocarcinoma.

- Ductal PCa previously called endometrial cancer is today considered a morphological variant of adenocarcinoma with two growth patterns: papillary and cribriform. The prognosis is rather poor (Bostwick, Kindrachuk et al. 1985).
- Mucinous cancer is almost always seen together with conventional PCa and is defined by tumor cells floating in extra-acinar mucin with at least 25% mucinous differentiation. There are controversial results concerning the biological behavior of this entity. While one study found that mucinous carcinoma rarely responds to hormonal therapy and often presents with advanced stage (Epstein and Lieberman 1985) another recent presented study on 47 radical prostatectomy specimens describes this variant as not more or even less aggressive than nonmucinous adenocarcinoma (Osunkoya, Nielsen et al. 2008).
- Signet-ring carcinoma is a clinically aggressive high-grade tumor (Alline and Cohen 1992).
- Small cell carcinomas (including the neuroendocrine carcinomas) are thought to arise from multipotential undifferentiated prostatic epithelium and have a poor prognosis, with a median survival of 7.7 months (Ro, Tetu et al. 1987).
- Transitional cell carcinomas arise from the periurethral glandular epithelium or from metaplastic prostatic epithelium. They are often aggressive tumors that do not respond to hormonal therapy (Young 2000).
- Lymphoepithelioma-like carcinomas are extremely rare. They are poorly differentiated with a syncytial growth pattern, prominent lymphocytic stroma and adverse clinical behavior (Young 2000).
- Biphasic tumors that contain an adenocarcinoma and a recognizable sarcomatous component such as carcinosarcomas are rapidly progressive (Lauwers, Schevchuk et al. 1993).
1.2.5 Prognostic and predictive factors in prostate cancer

Prediction of prognosis is one of the greatest challenges in tumor pathology (Burke, Bostwick et al. 2005). Despite attempts to introduce new biomarkers, histopathological grade remains the most important tissue-based predictor of prognosis for many cancer types in general and for PCa in particular. New markers with correlation to prognosis are described almost every month. Thus, this chapter can only be incomplete and is restricted to more or less established prognostic factors. Currently, there are no markers that predict response or resistance to a PCa specific therapy.

Prognostic factors, which predict relapse or progression independent of future treatment effects, can be stratified according to the College of American Pathologists (CAP) into three different categories (Bostwick, Grignon et al. 2000).

Category I: markers that are well supported by the literature and generally used in patient management. In this category preoperative PSA, TNM stage grouping, histologic grade (Gleason score) and surgical margin status are included. These prognostic markers are used in clinical practice in the form of nomograms for prediction of tumor progression (Kattan, Wheeler et al. 1999).

Category II: markers that are extensively studied biologically and/or clinically but with few clinical outcome studies. Factors included are tumor volume, histological type, and DNA ploidy.

Category III: markers that currently do not meet the criteria of category I or II. Factors in this category include perineural invasion, neuroendocrine differentiation, microvessel density (angiogenesis), nuclear roundness, chromatin texture, other karyometric factors, proliferation markers, prostate specific antigen derivatives, and other factors (oncogenes, tumor suppressor genes, apoptosis genes, etc).

There is a huge amount of new studies in the field of prognostic factors in PCa and borders between the above named categories are fluctuating. Hence, most novel biomarkers are not included in the CAP list.
1.2.5.1 The Gleason grading system

The Gleason score (GS) is today officially recommended by the World Health Organization as grading system for PCa (Eble JN 2004). A broad consensus is reached that PCa should be graded according to Gleason on needle biopsies, transurethral prostate resection (TURP) specimens and radical prostatectomy specimens (Bostwick, Grignon et al. 2000; Montironi, Lopez-Beltran et al. 2001; Montironi, van der Kwast et al. 2003; Eble JN 2004; Epstein, Amin et al. 2005; Montironi, Vela Navarrete et al. 2006; Srigley, Amin et al. 2006; Epstein, Srigley et al. 2007). Other grading systems like Mayo grading, National Prostatic Cancer Treatment Group (NPCTG) grading, NPCTG score, Böcking combined grading and Mostofi grading ("WHO grading") have never reached the same acceptance.

Donald F. Gleason developed his grading system from 1960 to 1974 while reviewing an accumulated number of about 5000 PCa patients in prospective randomized clinical trials at the Veterans Administration Cooperative Urological Research Group (VACURG) (Bailar, Mellinger et al. 1966; Gleason 1966; Mellinger, Gleason et al. 1967; Gleason and Mellinger 1974; Gleason 1977; Gleason 1988; Gleason 1992). The strength of the Gleason system lays in its testing on a large patient population with long follow-up and cancer-specific survival as an endpoint and its original presentation with an instructive schematic drawing (figure above).

Based solely on glandular architecture on hematoxylin and eosin stained prostatic tissue sections, the GS does not take cytological features into consideration. Extent of glandular differentiation and growth patterns are analyzed microscopically at relatively low magnification (2-10X lens magnification) resulting into five patterns illustrated in the original drawing above. The primary and secondary pattern, i.e. the most prevalent and the second most prevalent pattern are added to obtain a GS or sum (GS = Gleason
pattern 1 + Gleason pattern 2). The prevalence of patterns is determined by simple
visual estimation. The score can range from 2 to 10. If the tumor has only one pattern,
the GS is obtained by doubling that pattern. In daily practice a cut-off at 5% is often
used for inclusion in the GS. The Gleason system is described in detail and illustrated
in many publications (Humphrey 2004; Epstein, Allsbrook et al. 2005) and is also
available via the internet (http://pathology2.jhu.edu/gleason/).

The GS is probably the single most powerful predictor of patient outcome, which
correlates with multiple other parameters related to prognosis, such as age (Draisma,
Postma et al. 2006; Helpap in press; Helpap in press), serum PSA (Horninger, Rogatsch
et al. 1999; Helpap in press), clinical stage (Ramos, Carvalhal et al. 1999) and
pathological stage (Veltri, Miller et al. 1996; Helpap and Egevad 2006; Helpap in
press; Helpap in press). Numerous studies show that the GS is an independent and very
powerful prognostic factor, both for prediction of the natural history of PCa (Albertsen,
Fryback et al. 1995; Egevad, Granfors et al. 2002; Andren, Fall et al. 2006; Berney,
Fisher et al. 2007) and for assessment of the risk of recurrence after radical
prostatectomy (Epstein, Partin et al. 1996; Han, Partin et al. 2001; Hull, Rabbani et al.
2002; Han, Partin et al. 2003) or radiotherapy (Zagars, Ayala et al. 1995; Green,
Hanlon et al. 1998).

A weakness of the Gleason system is that it has a pronounced clustering in the mid
range of the scores. Some reports of radical prostatectomy series observed as many as
86% to 89% of the tumors presenting a GS 6 or 7 (Steinberg, Sauvageot et al. 1997;
Stamey, McNeal et al. 1999). Gleason pattern 1 is extremely rare. Pattern 2 is usually
mixed with some pattern 3 resulting in a GS 5. As a result GS 2, 3 and 4 are only
exceptionally assigned and it is recommended not to use scores 2-4 on needle biopsies
(Epstein 2000; Epstein, Allsbrook et al. 2005). Pattern 5 is usually mixed with some
pattern 4 resulting in a GS 9 and, therefore GS 10 is uncommon.

In the last ten years there has been a gradual shift in the grading of PCa (Epstein 2000;
Sengupta, Slezak et al. 2006) with a general trend towards upgrading. Albertsen et al.
discuss the clinical impact of these changes (Albertsen, Hanley et al. 2005). The
prognosis of GS 6 tumors will improve if most GS 5 tumors are moved into GS 6 while
the worst GS 6 cancers are moved into GS 7. This modifies the prognostic implications
of pathology reports, potentially leading to confusion in treatment decisions.
Comparisons with results from earlier studies will also be more difficult.

Despite these problems the GS remained mainly unchanged until 2005 when the
International Society of Urological Pathology (ISUP) provided new recommendations
for a modified Gleason grading (Epstein, Allsbrook et al. 2005). A modified diagram of the Gleason patterns was presented with changes mainly in patterns 3 and 4. (Table 2) GS reporting on core biopsies and prostatectomy specimens was also discussed. It was agreed that a GS should be assigned to each separate biopsy and that a global GS (a summary of all cancer present) could be optionally provided in the bottom line diagnosis. Furthermore, high-grade pattern of any quantity should be included in the GS of biopsies. For example a main tumor with most prevalent patterns 3 and 4 and a tertiary pattern 5 should be assigned a $3 + 5 = 8$ instead of $3 + 4 = 7$. For radical prostatectomy specimens where the entire tumor is available for examination the participants agreed that the GS of the main tumor should still include the primary and secondary patterns and a tertiary pattern of higher grade should only be mentioned in a separate comment in the pathology report. Other recent recommendations are in line with the ISUP recommendations (Eble JN 2004; Amin, Boccon-Gibod et al. 2005).

GS 7 was shown to be a separate prognostic category between GS 6 and 8 (Egevad, Granfors et al. 2002). The major increase of the likelihood of having adverse findings in the prostatectomy specimen or of the failure rate following prostatectomy or radiotherapy was from GS 6 to 7, indicating that the presence of Gleason pattern 4 cancer had a negative effect on prognosis. In line with this, multiple studies showed that GS 7 cancers with a primary pattern 4 have more advanced pathological stage and higher risk of disease progression after prostatectomy than those with a primary pattern 3 (Chan, Partin et al. 2000; Sakr, Tefilli et al. 2000; Makarov, Sanderson et al. 2002), while a single study failed to show any significant difference between GS 3+4 and 4+3 (Grober, Tsihlias et al. 2004). Consequently the roles of tertiary Gleason patterns 4 and 5 were investigated.

Tertiary Gleason patterns are defined as a grade that is higher than the primary and secondary patterns and therefore not included into the GS. It was shown that their presence had an adverse effect on prognosis (Stamey, Yemoto et al. 2000; Egevad, Granfors et al. 2002; Egevad, Granfors et al. 2002). For many years Gleason patterns comprising less than 5% of the tumor were not included in the GS. Pan et al. showed that tertiary high-grade components had an adverse impact on biological behavior even when small (<5%) (Pan, Potter et al. 2000). Men with any tertiary pattern 4 or 5 in a GS 5 or 6 tumor had a prognosis similar to those with a GS 7. Hattab et al. found that a tertiary Gleason pattern 5 was the strongest predictor of poor patient outcome in GS 7 prostate carcinomas, regardless of whether grade 3 or 4 was
Table 2: Summary of histological main features and changes in the Gleason grading proposed at the ISUP meeting 2005 (modified Gleason grading).

<table>
<thead>
<tr>
<th>ISUP modified Gleason grading</th>
<th>Histological main features</th>
<th>Loss of stromal differentiation</th>
<th>Dedifferentiation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Gleason grade 1b (core biopsies)</td>
<td>Exteroelastic glands with variable-sized nuclei, small to intermediate cell volume, and moderate cell density.</td>
<td>No</td>
<td>No</td>
</tr>
<tr>
<td>Gleason grade 2 (core biopsies)</td>
<td>Exteroelastic glands with variable-sized nuclei, small to intermediate cell volume, and moderate cell density.</td>
<td>No</td>
<td>Partial</td>
</tr>
<tr>
<td>Gleason grade 3 (core biopsies)</td>
<td>Exteroelastic glands, small to intermediate cell volume, and moderate cell density.</td>
<td>No</td>
<td>Total</td>
</tr>
<tr>
<td>Gleason grade 4 (core biopsies)</td>
<td>Exteroelastic glands, small to intermediate cell volume, and moderate cell density.</td>
<td>Partial</td>
<td>Total</td>
</tr>
<tr>
<td>Gleason grade 5 (core biopsies)</td>
<td>Exteroelastic glands, small to intermediate cell volume, and moderate cell density.</td>
<td>No</td>
<td>Total</td>
</tr>
</tbody>
</table>

Essentially no glandular differentiation, composed of solid sheets, cords, or single cells. Epithelial cell nuclei with central nuclei.
the primary grade (Hattab, Koch et al. 2006). Patel et al. documented that GS 7 tumors diagnosed on biopsies with tertiary pattern 5 had a similar prognosis as GS 8 tumors when treated by radiotherapy or radical prostatectomy (Patel, Chen et al. 2007).

These findings opened a discussion on whether the Gleason system needed to be modified to account for the prognostic information provided by high-grade patterns (Stamey, McNeal et al. 1999; Pan, Potter et al. 2000; Vis, Roemeling et al. 2007). In the late 1990s, the Stanford group described the prognostic impact of the percentage of Gleason grade 4 or 5 (%GG4/5), i.e. the proportion of the tumor occupied by high-grade cancer. They were able to show that it might provide prognostic information that were independent from that given by the conventional Gleason system and proposed that the conventional Gleason system should be abandoned in favor of %GG4/5 (Stamey, McNeal et al. 1999). They showed previously that %GG4/5 predicted lymph node metastases after radical prostatectomy (McNeal, Villers et al. 1990; Stamey, McNeal et al. 1999; Stamey, Yemoto et al. 2000). Others were able to confirm that %GG4/5 was superior to the Gleason grading system in predicting recurrence (Cheng, Koch et al. 2005; Vis, Roemeling et al. 2007). In a Swedish study of TURP detected PCa’s managed by watchful waiting, %GG4/5 was a better predictor of disease-specific survival than the GS and confirmed that %GG4/5 conveyed prognostic information that was independent of that given by the GS (Egevad, Granfors et al. 2002). There are conflicting data about the correlation between %GG4/5 in prostate biopsies and radical prostatectomy specimens. In two studies, Stamey et al. observed a rather strong correlation (r²=0.63 and 0.57, respectively) (Stamey 1995; Stamey, McNeal et al. 1999) while Rubin et al. found this correlation to be weaker (r²=0.32) (Rubin, Mucci et al. 2001). The latter authors also indicated a low sensitivity (38%) but a high specificity (96%) for predicting any pattern 4/5 carcinoma in prostatectomy specimens.

%GG4/5 did not gain general acceptance for grading of PCa despite its excellent ability to predict prognosis and whether it will be used in clinical practice remains to be seen.

Pan et al. proposed to renounce at the GS in favor of a modified GS (mGS), which was defined as the sum of the primary Gleason grade plus the worst grade, even if less than 5% of the tumor (Pan, Potter et al. 2000). The authors emphasized that further studies were needed to confirm their findings, but only few studies had specifically addressed the mGS.

In a recent study Helpap et al. demonstrated that the mGS minimized undergrading of prostatic carcinomas in biopsies and improved the agreement between biopsies and
radical prostatectomy specimens. The overall exact agreement between GS of needle biopsies and radical prostatectomy specimens improved from 58 to 72% with modified Gleason grading (Helpap and Egevad 2006; Helpap in press). In addition, it was later found that mGS correlated with age, serum PSA, percent positive biopsies and percent cancer length in prostate biopsies (Helpap in press). In their consecutive needle biopsy series from the years 1995, 2000, 2006 and 2007 there was a stage shift downwards as indicated by lower serum PSA and percent positive biopsies. Yet the tumor grades were higher in 2006 and 2007 when mGS was used.

Veloso et al. investigated the interobserver agreement and prediction of GS and mGS in needle biopsy and in surgical specimen of PCa (Veloso, Lima et al. 2007). The authors found that the mGS was not superior to GS neither for the agreement between the biopsies and prostatectomy specimen nor in interobserver reproducibility. However, this study had a limited number of cases with a tertiary pattern on biopsy (0%, 2%, 8% amongst the three observers), making it difficult to show differences between conventional GS and mGS.

After the new ISUP recommendations to use mGS for grading of needle biopsies it will be necessary to re-iterate some previous studies performed with the conventional GS, including its prognostic value, interobserver reproducibility and correlations with other clinicopathological parameters.

1.2.5.1.1 Prediction of Gleason score by core needle biopsies

The pronounced intratumoral grade heterogeneity of PCa raises concern about the ability of biopsies to predict prostatectomy grade. Over 50% of radical prostatectomy specimens contain cancer of at least three different Gleason patterns (Aihara, Wheeler et al. 1994) and cancer of a single grade was present in only 16% of the specimens in one study (Ruijter, van de Kaa et al. 1996). Of individual tumor foci, 58% have a single grade, but most of these foci are very small (Ruijter, van de Kaa et al. 1996). In a study by Arora et al, two or more cancer foci were present in 87% of all radical prostatectomy specimens (Arora, Koch et al. 2004). In only 9% of the cases with multifocal cancer, all tumor foci had Gleason grades that were the same as the overall Gleason grades of the cases.

Several studies investigated the ability to predict the prostatectomy GS by core needle biopsies (Spires, Cibull et al. 1994; Cookson, Fleschner et al. 1997; Steinberg,
Sauvageot et al. 1997; Carlson, Calvanese et al. 1998; Djavan, Kadesky et al. 1998; Egevad 2001; Gregori, Vieweg et al. 2001; Humphrey 2003). In a compilation of data on 3,789 patients from 18 studies, exact correlation of GSs was found in 43% of cases and correlation plus or minus one Gleason core unit in 77% of cases (Humphrey 2003), while Egevad et al. found correlation within one GS unit in more than 90% of the cases when an average of 10 biopsies were taken (Egevad, Norlen et al. 2001). In a recent review there was an exact correlation in 28.2 - 67.9% of the cases (Mazzucchelli, Barbisan et al. 2005). Undergrading of PCa in needle biopsy is the most common problem with a frequency of 24.5 to 60%, while overgrading occurs only in 5.2 to 32.3% of the cases (Gleason 1996; Djavan, Kadesky et al. 1998; Grossfeld, Chang et al. 2001; Humphrey 2003).

Agreement might be explained to some part by the high proportion of tumors assigned as GS 6 or 7 in needle biopsies and prostatectomy specimens. However, with wk statistics, the distribution is taken into account. Weighted kappa values between 0.41 and 0.50 were obtained indicating a moderate agreement (Rubin, Dunn et al. 2000; Egevad, Norlen et al. 2001).

There are several sources of discrepancy between biopsy vs. prostatectomy GS. Sampling errors relating to the small amount of tissue removed by core needle biopsies are perhaps the most important factor. The most common type of sampling error occurs when there is a higher grade component present in the radical prostatectomy specimen, which is not sampled on needle biopsy, thus leading to undergrading (Steinberg, Sauvageot et al. 1997). For example, a prostate biopsy with a GS 3 + 3 = 6 could have missed a grade 4 component, resulting in a postoperative GS 3 + 4 = 7. Another potential problem are so called minimal cancer foci in biopsies that might not be representative of the entire tumor. However, surprisingly the grade concordance between biopsy and prostatectomy in these cases is equivalent or only slightly worse than in larger cancers (Steinberg, Sauvageot et al. 1997; Rubin, Dunn et al. 2000; Egevad, Norlen et al. 2001). Overgrading may result when a high-grade pattern is selectively represented in needle biopsies while this pattern is only a minor element in the prostatectomy specimen and, hence, should not be included in the GS.

Borderline cases at the interface between two Gleason patterns may cause reproducibility problems and therefore jeopardize biopsy prediction of prostatectomy GS (Steinberg, Sauvageot et al. 1997; Rubin, Dunn et al. 2000; Egevad, Norlen et al. 2001). Some pathologists tend to underage needle biopsies, e.g. by ignoring cancer infiltration between benign glands, which precludes a Gleason pattern 2. Small areas
of gland fusion, cribriform patterns and infiltrative growth representing a high-grade tumor are also easily overlooked (Steinberg, Sauvageot et al. 1997; Rubin, Dunn et al. 2000; Egevad, Norlen et al. 2001).

Iatrogenic morphological changes in prostate samples including the effects of hormonal treatment and radiotherapy or crush artifacts caused by surgical procedures or specimen handling may lead to deceptively high Gleason grade. Therefore, it is recommended not to grade tumor tissue showing such changes.

1.2.5.1.2 Reproducibility of the conventional Gleason score

Grading of PCa is particularly difficult because of the pronounced morphological heterogeneity of this tumor, raising concerns about the interobserver reproducibility of the Gleason system. Several studies investigated the level of interobserver variability, reviewed by Allsbrook et al. (Allsbrook, Mangold et al. 2001). Pathologists who were inexperienced in the Gleason system tended to undergrade (Steinberg, Sauvageot et al. 1997; Allsbrook, Mangold et al. 2001; Egevad 2001). The vast majority of tumors graded as GS 2 to 4 on core biopsy were scored as GS 5 to 6 or higher when reviewed by experts in urological pathology (Allsbrook, Mangold et al. 2001). GS 2 to 4 on core biopsy had poor reproducibility even among urological pathology experts and should therefore be avoided (Epstein 2000).

For many reasons, the results of these studies are difficult to compare. The number and the type of specimens differed. Some studies were preceded by a tutorial. In other studies the specimens were selected rather than consecutive. Different number of observers also makes comparisons difficult. Svanholm et al. reached a wk for GS of 0.70 but only two observers were involved, which may overestimate the agreement compared with a larger group of observers that have not been trained together (Svanholm and Mygind 1985). Lessells et al. had 12 observers and a wk of 0.45 (Lessells, Burnett et al. 1997). In a study by Allsbrook et al. an overall mean wk of 0.44 (range 0.00 to 0.88) was reached with 10 general pathologists and a set of 38 selected biopsy specimens (Allsbrook, Mangold et al. 2001). Interobserver agreement among urological pathologists described by the same authors was moderate to substantial (wk 0.56 to 0.70) (Allsbrook, Mangold et al. 2001). A British group of pathologists reached a wk 0.54 for GS groups 2-4, 5-6, 7, 8-10 (Melia, Moseley et al. 2006). Despite all these difficulties and disagreements, improvement is possible. It was shown that the interobserver reproducibility was strongly correlated with course paticipation for
Gleason grading at a meeting or a course (Allsbrook, Mangold et al. 2001) and that it could be significantly improved by formal educational efforts (Kronz, Silberman et al. 2000; Egevad 2001).

1.2.5.2 Other morphology-based factors

1.2.5.2.1 Pathological stage

The TNM system for histopathological staging of PCa has become the predominant method for assessing the extent of disease and includes several parameters (Sobin 2002). Stage determining factors such as extra-prostatic extension (EPE), seminal vesicle invasion (SVI) and lymph node metastasis are described in the following sections.

1.2.5.2.2 Extra-prostatic extension (EPE)

EPE is common in prostatic adenocarcinoma with a frequency between 30-50% (Davis, Pisansky et al. 1999; Gilliland, Hoffman et al. 1999). In current prostatectomy series the incidence of EPE shows declining numbers, maybe due to the result of PSA screening and the consequent stage migration. In one study EPE in radical prostatectomy specimens decreased from 81% to 36% from 1987 to 1997 (Jhaveri, Klein et al. 1999). The extent of EPE may have prognostic importance. Patients with focal EPE were reported to have an intermediate risk of recurrence between those with organ-confined disease and those with established EPE (Epstein, Pizov et al. 1993). The association of a high-grade tertiary Gleason component and organ-confined PCa versus focal or extensive EPE was reported (Pan, Potter et al. 2000).

1.2.5.2.3 Seminal vesicle invasion (SVI)

SVI by PCa is detected in less than 10% of patients (Bloom, Richie et al. 2004) and is associated with high tumor grade, large tumor volume, EPE, lymph node metastasis, and poor prognosis (Ohori, Scardino et al. 1993). SVI in radical prostatectomy specimen markedly diminishes the likelihood of cure. In contemporary series 5-year biochemical progression-free rates range from 5% to 60% (mean 34%) (Epstein, Pizov et al. 1993; Trapasso, deKernion et al. 1994; D'Amico, Whittington et al. 1995; Ohori, Wheeler et al. 1995; Catalona and Smith 1998; Debras, Guillonneau et al. 1998; Tefilli,
SVI is associated with high PSA failure rates after radical prostatectomy, and subsequent distant metastases (Bloom, Richie et al. 2004). The results of Freedland et al. (Freedland, Aronson et al. 2004) revealed that patients with SVI had higher preoperative PSA levels, higher clinical stage, higher tumor grade and were more likely to have concomitant EPE or a positive surgical margin. A significantly higher frequency of tertiary Gleason grades was observed in patients with SVI (Pan, Potter et al. 2000; van Oort, Schout et al. 2005). In SVI, cancer is seen penetrating the muscular wall of the extraprostatic portion of the seminal vesicle (Potter, Epstein et al. 2000). In contrast, tumors invading the intraprostatic part of the seminal vesicle apparently do not have a worse prognosis than organ-confined tumors (Epstein, Amin et al. 2005).

1.2.5.2.4 Positive pelvic lymph nodes

The adverse prognosis of PCa with metastases to pelvic lymph nodes is universally accepted. Nodal metastases are associated with tumor progression in between 8.5% and 40% of cases in some earlier reports (Fowler and Whitmore 1981; Epstein, Amin et al. 2005). The incidence of positive pelvic lymph nodes decreased recently due to earlier detection of PCa and better patient selection and it was only seen in 1.2% of cases in 1999 at the Johns Hopkins Hospital (Epstein, Amin et al. 2005). It is controversial whether other factors can stratify patients with lymph node metastases in different prognostic groups. There are conflicting data on the prognostic impact of number, size and volume of the metastases as well as extra-nodal extension (Steinberg, Epstein et al. 1990; Sgrignoli, Walsh et al. 1994; Griebling, Ozkutlu et al. 1997; Cheng, Bergstralh et al. 1998; Potter, Mangold et al. 2000).

1.2.5.2.5 Positive resection margins (PRM)

The overall incidence of margin positivity ranges from 20-30% and is more frequent in large volume, high Gleason grade tumors (Epstein, Pizov et al. 1993; Cheng, Darson et al. 1999; Emerson, Koch et al. 2005). Residual tumor cells in patients with PRM might lead to relapse and though significantly affect disease outcome (Epstein, Pizov et al. 1993). Patients with PRM have a 50-60% 5-year risk of disease progression and a significant higher risk of PSA recurrence (Connolly, O'Toole et al. 2004). A tertiary Gleason grade was strongly associated with the incidence of PRM (Mosse, Magi-
Galluzzi et al. 2004; van Oort, Schout et al. 2005). PRM were an independent prognostic factor in two large studies (n = 1389) (Swindle, Eastham et al. 2005) and (n = 2518) (Blute, Bergstralh et al. 2001).

1.2.5.2.6 Perineural invasion

Perineural invasion, i.e. migration of tumor along the perineural sheath, is recognized as a low-resistance pathway of tumor penetration into the extra-prostatic space (Bastacky, Walsh et al. 1993). Perineural invasion on needle biopsy has been reported to be a specific marker for EPE of the tumor in a prostatectomy specimen (Bastacky, Walsh et al. 1993). Thus it should be named in the pathology report on prostate biopsies. The prognostic significance of perineural invasion remains controversial. A recent study found independent significance only when the percentage of tumor on the needle biopsy cores was not considered (Rubin, Bassily et al. 2000). In several studies perineural invasion did not predict tumor progression (van den Ouden, Hop et al. 1997; Maru, Ohori et al. 2001; Ito, Nakashima et al. 2003).

1.2.5.2.7 Vascular invasion

Vascular invasion in radical prostatectomy was reported in several studies with as much as 12% to 46% of cases (Bahnson, Dresner et al. 1989; van den Ouden, Hop et al. 1997; de la Taille, Rubin et al. 2000). McNeal et al. showed that vascular invasion was an independent prognostic factor in tumors with a volume of 4 to 8 cm$^3$ but not in tumors <4 cm$^3$ (McNeal and Yemoto 1996). Others demonstrated vascular invasion to be an independent prognostic factor (Bahnson, Dresner et al. 1989; Herman, Wilcox et al. 2000). The CAP recommends reporting vascular invasion in prostatectomy specimens when present.

1.2.5.2.8 Tumor volume

Tumor volume is a significant predictor of pathologic stage, lymph node and distant metastasis, and overall disease outcome (McNeal, Villers et al. 1990; McNeal 1992; Stamey, Freiha et al. 1993). As described by McNeal et al. loss of differentiation and metastatic potential were strongly correlated with tumor volume (McNeal, Bostwick et al. 1986). However, it is controversial whether tumor volume is an independent
prognostic factor. Epstein et al. showed that Gleason grade, surgical margins, and tumor volume – independently from each other - were strongly correlated with progression in univariate regression analysis. However, in multiple regression analysis, tumor volume did not provide independent prognostic information beyond that provided by GS and margin status (Epstein, Carmichael et al. 1993). Similarly, in a study on 1302 cases, Kikuchi et al. did not find tumor volume to be an independent prognostic factor (Kikuchi, Scardino et al. 2004). The clinical importance of tumor volume has probably decreased in recent years because of stage migration. A high proportion of cancers are nowadays small when diagnosed and their volume is then less likely to discriminate between prognostic categories. Therefore, tumor volume is not measured routinely in most laboratories (Egevad in press). The number and length of involvement of multiple (sextant or octant) needle biopsy cores have been shown to predict overall tumor volume, pathologic stage, and disease outcome (Ravery, Chastang et al. 2000).

1.2.5.2.9 Nuclear morphometry (kariometry)

In 1982 Diamond et al. introduced nuclear morphometry for prediction of prognosis of PCa (Diamond, Berry et al. 1982). Since then many histomorphological studies have been performed on the use of automated analyzing techniques (Epstein, Berry et al. 1984; Clark, Askin et al. 1987; Eichenberger, Mihatsch et al. 1987; Mohler, Partin et al. 1988; Partin, Walsh et al. 1989). While earlier studies focused on nuclear size and shape, more recent studies analyzed the distribution and pattern of nuclear chromatin. Indolent well differentiated PCa in general shows almost perfectly rounded nuclei while irregular nuclei are associated with high tumor grade, metastatic potential and decreased survival. Eichenberger and colleagues found that elliptical shape measurement was the best morphometrical measure for distinction between PCa with good and poor prognosis (Eichenberger, Mihatsch et al. 1987). Partin et al. developed the Hopkin’s Morphometry System for morphometric analyses of 15 different shape descriptors by 17 different statistical tests (Partin, Walsh et al. 1989). They found that the elliptical shape descriptor showed the best ability to discriminate PCa patients in different prognostic groups. A prognostic score including stage, GS, age and variance of nuclear roundness stratified patients in three prognostic groups. Veltri et al. developed the quantitative nuclear grade (QNG) out of 11 multivariate significant nuclear morphometric descriptors (NMD) (Veltri, Miller et al. 1996). Biochemical
recurrence was predicted by QNG with a high specificity and sensitivity. NMD and DNA ploidy could be measured on the same Feulgen stained sections, which is time and cost effective.

1.2.5.2.10 Microvessel density

Increased vascularity has been recognized in a variety of tumors. In PCa, angiogenesis was associated with adverse outcome, pathological stage, and presence of metastasis (Weidner, Carroll et al. 1993). Microvessel density correlated with disease progression after radical prostatectomy in some studies (Silberman, Partin et al. 1997; Strohmeyer, Rossing et al. 2000) but not in others (Gettman, Bergstralh et al. 1998). These conflicting results may reflect methodological differences. Microvessel density can be measured in “hot spots” or in random areas, which may influence results as PCa typically shows increased angiogenesis in the center of the tumor. Variable sensitivity of different endothelial markers (CD 31, CD 34, Factor VIII) contributes to difficulties in comparing different studies.

1.2.5.3 Biomarkers

For decades morphology assessment of the specimens was the only tool pathologists had to evaluate PCa tissue. Due to new approaches, a large amount of additional biological data is gathered from the DNA, mRNA and protein level, thus opening the exciting field of molecular pathology. The use of biomarkers is one of the aspects (Flaig, Nordeen et al. 2007).

Biomarkers are defined as measurable indicator of a specific biological state that can serve for many clinical purposes: to screen for, diagnose or monitor a disease, to guide molecularly targeted therapy or assess therapeutic response. Molecular markers comprise several entities as their discovery uses different strategies from cDNA microarray studies to proteomics analysis.

After many years of sustained research efforts we can talk today about a coherent and comprehensive biomarker pipeline that in general lines includes: candidate biomarker discovery, verification, assay optimization and validation (Rifai, Gillette et al. 2006). Validation of candidate tissue biomarkers is a step of paramount importance and for this we used the TMA method in our studies. First described by Kononen et al. (Kononen, Bubendorf et al. 1998) TMA plays now an exceedingly important roll in
high-throughput validation (Dhanasekaran, Barrette et al. 2001). The TMA technology is applicable for in situ analysis of target structures on protein (immunohistochemistry), RNA (RNA-in-situ-hybridization) and DNA (fluorescence-in-situ-hybridization) level. Kallionemi et al. described the possibility to place 1000 samples within a 45 x 20 mm² area when 0.6 mm punches are utilized and cores are deposited with 0.8 mm spaces (Kallioniemi, Wagner et al. 2001). An example is the large-scale validation of human proteins by TMA in the "Human Proteome Resource” project (http://www.proteinatlas.org/).

To construct a TMA, small core biopsies are taken from morphologically representative areas of paraffin-embedded tumor tissues and assembled on a recipient paraffin block in a grid manner (see figure). The method is described in detail in several articles and reviews (Kononen, Bubendorf et al. 1998).

Extensive planning is necessary before a TMA is built. A. Kajdacsy-Balla gave a contemporary overview. (Kajdacsy-Balla, Geynisman et al. 2007)

A commonly discussed problem in TMA construction is whether a tumor can be represented in TMA cores, which usually have a diameter between 0.6 mm and 1 mm. This is a particularly relevant discussion for a heterogeneous tumor like PCa. Rubin et al. addressed this issue by analyzing Ki-67 expression in PCa, comparing TMA with conventional immunohistochemistry slides (Rubin, Dunn et al. 2002). They concluded that 3 to 4 TMA cores are optimal for evaluating PCa biomarkers.

Loss of tissue during sectioning and staining of TMAs is a common problem. In previous studies it was shown that 10 to 30% of tissue might be lost during the entire TMA process (Bubendorf, Kononen et al. 1999; Schraml, Kononen et al. 1999; Mucci, Akdas et al. 2000; Richter, Wagner et al. 2000; Hoos and Cordon-Cardo 2001). To face this problem Kramer et al. recommends the use of positively charged slides, covered with an adjuvant adherent coating (Superfrost®-PLUS, Menzel, Braunschweig, Germany) (Kramer, Merseburger et al. 2007). They further described that the stability of cores of 1 mm in diameter was superior to that of 0.6 mm cores during sectioning.
and staining. This recommendation is also supported by the Human Proteome Resource project (personal communication). The use of an adhesive tape system has been regarded as controversial. Henshall et al. recommended the use of adhesive tape (Henshall 2003), while Packeisen et al. and Hoos et al. raised concerns of false negative results (Hoos and Cordon-Cardo 2001; Packeisen, Korsching et al. 2003).

There are many studies that evaluated either individual biomarkers or a group of biomarkers belonging to a single family based on structure or function in the diagnosis of PCa as well as for risk stratification and prognosis.

1.2.5.3.1 Prostate specific antigen (PSA)

Serum PSA is the most widely used biomarker for PCa screening and detection, contributing to patient stratification into different prognostic categories (Partin, Mangold et al. 2001). PSA is also the major mean of monitoring patients for tumor recurrence following treatment. But PSA testing is associated with several limitations as for example the absence of a PSA cut point value with simultaneous high sensitivity and high specificity for monitoring healthy men for PCa. The relationship between serum PSA and patient outcome is influenced by coexisting prostatic hyperplasia (Beduschi and Oesterling 1997), patient age (Richardson and Oesterling 1997), and GS (Partin, Carter et al. 1990). The fact that PSA demonstrates a PCa risk at all values raises many challenges for the clinicians (Thompson, Ankerst et al. 2005).

Many efforts were invested in improving the operating characteristics of PSA like PSA velocity, PSA density, PSA doubling time, age specific PSA ranges, unbound PSA, percent complexed PSA and PSA isoforms. None of them brought profound improvements (Parekh, Ankerst et al. 2007).

1.2.5.3.2 DNA ploidy

The first report on the correlation between DNA ploidy and prognosis of PCa appeared in 1966 (Tavares, Costa et al. 1966). The majority of retrospective studies showed that aneuploid DNA content in PCa independently predicted poor prognosis (Winkler, Rainwater et al. 1988; Montgomery, Nativ et al. 1990; Peters, Miles et al. 1990; Foster, McLoughlin et al. 1992; Peters-Gee, Miles et al. 1992; Ross, Nazeer et al. 1993). Forsslund and Zetterberg studied the prognostic role of DNA content in patients with
long-term follow-up (Forsslund and Zetterberg 1990). Patients who died within 3 years of diagnosis consistently had DNA stem lines at 3c and 6c, whereas long-term survivors (> 15 years) had stem lines at 2c and 4c. Stephenson et al. found a strong correlation between ploidy status of pelvic lymph node metastases and mean survival time (Stephenson, James et al. 1987). In the Mayo Clinic prostatectomy series, ploidy was a significant predictive factor in multivariate analysis (Ward, Slezak et al. 2005). Beside serum PSA, DNA ploidy is the only biomarker of PCa that has been recommended for clinical use by the CAP.

1.2.5.3.3 Other biomarkers

Plenty of individual biomarkers were tested in different biospecimens such as serum, plasma, prostate tissue, seminal fluid, and urine with the purpose to either replace or augment the existing use of PSA. They are too many to be able to provide a comprehensible list. A detailed account of these biomarkers is beyond the scope of this thesis. An overview is provided by the following reviews (Bradford, Tomlins et al. 2006; Flaig, Nordeen et al. 2007; Parekh, Ankerst et al. 2007).

1.2.5.4 Potential prognostic factors in prostate cancer of interest for this thesis

1.2.5.4.1 Pancreatic duodenal homebox (PDX-1)

The pancreatic duodenal homeobox-1 gene (PDX-1) is a transcription factor suggested to be a master regulator of both pancreatic development and of pancreatic islet cell differentiation, especially the β-cell phenotype. It is also identified and described under its synonyms IPF-1, STF-1 and IDX-1.

During embryonic development, PDX-1 involvement is necessary for differentiation of exocrine and endocrine pancreas (Ashizawa, Brunicardi et al. 2004). In adults, PDX-1 serves to maintain islet cell physiology by activating gene transcription of insulin, somatostatin, islet amyloid polypeptide, glucose transporter type 2 and glucokinase (Ashizawa, Brunicardi et al. 2004).

Many studies underline the importance of the PDX-1 transcription factor in pancreas development. Reduced PDX-1 levels in mice lead to insulin deficiency, hyperglycemia and increased beta cell apoptosis (Ahlgren, Jonsson et al. 1998; Brissova, Shiota et al. 2002; Johnson, Ahmed et al. 2003). Targeted disruption of the PDX-1 gene in mice

Up to date PDX-1 is a protein primarily connected with different pancreatic disorders, e.g. diabetes mellitus, insulinoma and pancreatic cancer (Habener, Kemp et al. 2005; Wang, Li et al. 2005). Furthermore Koizumi et al. describe that PDX-1 expression in pancreatic cancer correlates with histological grade and lymph node metastasis and was also an independent prognostic factor (Koizumi, Doi et al. 2003). PDX-1 expression or overexpression is reported in several human cancers such as gastric (Leys, Nomura et al. 2006), breast, colon, prostate and kidney cancers (Wang, Li et al. 2005). Noteworthy in some cases are the elevated levels of PDX-1 in both malignant and adjacent benign tissues specimens from the same patients, whereas non-malignant controls do not express PDX-1 (Wang, Li et al. 2005). These data suggest that PDX-1 might be of interest as an early marker in carcinogenesis and to our knowledge no studies investigated the PDX-1 expression in precursor lesions of PCa or its correlation with prognosis of invasive PCa.

1.2.5.4.2 Heat shock proteins 27, 60 and 70

Heat shock proteins (HSPs) are considered useful as diagnostic or prognostic predictive factors in a variety of tumors, reviewed by Ciocca et al. (Ciocca and Calderwood 2005). HSPs are highly conserved, ubiquitous molecules required for proper folding, localization and stability of cellular proteins, as well as degradation of senescent proteins. They also act as molecular chaperones, protecting cells against stress-related injury (Fuller, Issels et al. 1994). HSPs are beneficial to the normal cell, but as stress conditions are commonplace in tumors, cancer cells can also use HSPs in response to stress, leading to increased expression (Gibbons, Watson et al. 2000). HSP expression levels are altered in a wide range of human neoplasms, e.g. carcinomas of the prostate, breast, female genital tract, gastrointestinal tract, liver, pancreas, lung, skin and urinary tract, hematological malignancies, sarcomas and endocrine adenomas (Ciocca and Calderwood 2005). In breast cancer, it is shown that increased expression of HSP27 and HSP70 are useful prognostic factors that correlate with resistance to
chemotherapy, significantly shorter disease-free survival, increased cell proliferation, poor differentiation and lymph node metastases (Ciocca, Clark et al. 1993; Lazaris, Chatzigianni et al. 1997; Vargas-Roig, Fanelli et al. 1997; Vargas-Roig, Gago et al. 1998).

It is suggested that HSPs might also be of interest as prognostic markers for PCa (Cornford, Dodson et al. 2000; Lebret, Watson et al. 2003; Kurahashi, Miyake et al. 2007). Other studies show that certain HSPs inhibited apoptosis and may therefore serve as independent survival factors, especially in androgen independent PCa (Thomas, Brown et al. 1996; Bostwick 2000; Gibbons, Watson et al. 2000).

Our research group recently demonstrated that the protein profile of PCa differed from that of benign prostatic tissue (Lexander, Palmberg et al. 2005). By 2-dimensional gel electrophoresis and mass spectrometry, several proteins were shown to be overexpressed in PCa, among them HSP60 and HSP70. Subsequently, we found correlation between HSP60 and HSP70 overexpression and prognostic factors such as Gleason score and DNA ploidy (Lexander, Palmberg et al. 2006).
2 AIMS OF THE STUDY

• To investigate the reproducibility of conventional Gleason score, percent Gleason grade 4/5 and modified Gleason score in radical prostatectomy specimens and prostate biopsies.

• To evaluate whether percent Gleason grade 4/5 in radical prostatectomy specimens can be predicted by multiple core biopsies.

• To analyze the prognostic value of PDX-1 and HSP 27, 60 and 70 in prostate cancer.

• To evaluate the role of PDX-1 in prostate cancer pathogenesis.
3 MATERIAL AND METHODS

3.1 REPRODUCIBILITY AND PREDICTION STUDIES (PAPERS I TO IV)

**Paper I and III:** A consecutive series of 69 radical prostatectomy specimens was collected at the Karolinska University Hospital, Solna, Sweden, from January to December 2000. The clinical stages T1b, T1c, T2 and T3 were diagnosed in 3 (4.3%), 48 (69.6%), 17 (24.6%), and 1 (1.4%) men, respectively. Mean preoperative serum PSA was 10.2 ng/ml (range 2.5 to 44). SVI, EPE and positive margins were present in 13 (19%), 41 (59%), and 33 (48%) cases, respectively. None of the patients had received hormonal treatment or radiotherapy prior to prostate biopsy. The prostates were formalin fixed overnight, inked and sliced horizontally at 4 mm intervals. The slices were cut in 2-6 segments and the prostate was totally embedded. A single hematoxylin and eosin section from each PCa was circulated among 4 genitourinary pathologists for a total of 276 responses and 414 pair wise comparisons. No consensus training preceded the study. Sections were reviewed. GS, %GG4/5 and mGS were calculated. Results were compared in pairs and wk values calculated for GS, %GG4/5 and mGS.

**Paper II:** A consecutive series of needle biopsies from 69 men diagnosed with PCa was collected. The biopsies were taken according to a standardized octant protocol and received between February and June 2001 at the Department of Pathology and Cytology at the Karolinska University Hospital, Solna, Sweden. The mean age at diagnosis was 67 years (range 50 to 88). None of the patients had received hormonal treatment or radiotherapy prior to prostate biopsy. The biopsies were formalin fixed overnight in separate containers and subsequently dehydrated, paraffin embedded in separate blocks, cut at 4 µm and stained with hematoxylin and eosin. Areas with cancer were marked on the slides with India ink along each biopsy. Tumor length was measured with a ruler. Average number of biopsies positive for cancer and total cancer length were 4 (range 1 to 9) and 28 mm (range 0.5 to 141 mm), respectively. In 15 men, cancer was present in a single biopsy, and in 22 men, at least 6 biopsies contained cancer. A total of 279 slides containing PCa were circulated among the same four pathologists as in Papers I and III. No consensus training for the estimation of %GG4/5 preceded the study. Slides were reviewed and GS and %GG4/5 were assessed for each
case. To avoid any bias, results from our recent study on reproducibility of %GG4/5 in prostatectomy specimens were not evaluated until all slides of the current study had been reviewed. The %GG4/5 was estimated as 0%, focal (5% or less) and subsequently at 10% intervals (11-20%, 21-30% etc.).

**Paper IV:** A consecutive series of radical prostatectomy specimens and preoperative needle biopsies obtained from 121 patients at Uppsala University Hospital, Sweden from November 1993 to March 1999 were considered for the study. After exclusion of six men with cancer detected by transurethral resection (one stage T1a and five T1b), 115 remained for inclusion in the study. Clinical stage was T1c, T2, and T3 in 47 (40.9%), 67 (58.3%), and 1 (0.9%) men, respectively. The mean preoperative PSA serum level was 12.9 ng/ml (median 10, range 2-61). Median age at surgery was 63 years (range 40 to 73). The biopsies and the prostatectomy specimens were handled as described in Papers I – III, except that the prostatectomy specimens were whole-mounted. All biopsy and prostatectomy specimens were reviewed by one pathologist while unaware of their identities. The GS and %GG4/5 of the main tumors were assessed. The combined %GG4/5 of all cancer foci was calculated by weighting the %GG4/5 of each focus against their volumes.

**3.2 PROTEIN EXPRESSION STUDIES (PAPERS V AND VI)**

### 3.2.1 Material

**Paper V and VI:** A consecutive series of 289 radical prostatectomy specimens was collected from 1998 to 2002 at Karolinska University Hospital, Solna, Sweden. The mean age of the patients at surgery was 61.1 years (median 61.3, range 46.2-74.1). Mean preoperative serum PSA was 9.0 ng/ml (median 7.7, range 0.5-58). The clinical stage was T1c in 191 (66.1%), T2 in 94 (32.5%), and T3 in 4 (1.4%) men. The GS of the tumors in the prostatectomy specimens was distributed as follows: GS 5: 10 (3.5%), GS 6: 127 (43.9%), GS 7: 118 (40.8%), GS 8: 14 (4.8%), GS 9: 19 (6.6%) and GS 10: 1 (0.3%). The Gleason patterns of individual TMA cores were 2 in 16 (1.8%), 3 in 597 (68.9%), 4 in 224 (25.8%), and 5 in 30 (3.5%) cores. The tumors showed EPE, positive surgical margins and SVI in 135 (46.7%), 133 (46.0%), and 36 cases (12.5%), respectively. The study also included 40 radical prostatectomy specimens from 2005. The GS of these cancers were 5 in 1 (2.5%), 6 in 18 (45%), 7 in 20 (50%), and 9 in 1
(2.5%) case. A consecutive series of 20 lymph nodes with PCa metastases from 2002 to 2005 was collected. None of the patients had received hormonal therapy or radiotherapy prior to surgery.

3.2.2 Tissue microarray (TMA) construction

Four different TMAs were constructed, including material from the Karolinska University Hospital, Solna, Sweden:

1. A prognostic TMA including 289 consecutive radical prostatectomy specimens.
2. A pathogenesis TMA from 40 radical prostatectomy specimens containing benign tissues (normal PZ, benign hyperplasia from TZ), atrophy, HGPIN and invasive carcinoma from the same cases.
3. A metastasis TMA that contained 20 lymph node metastases of PCa from 2002 to 2005.
4. Test TMAs, including PCa’s with different GS, PCa cell lines and benign prostatic tissue for antibody titration and pilot studies.

A Beecher Manual Tissue Arrayer 1 with 1 mm core punches was used for all TMAs. For the prognostic TMA we constructed 14 TMA blocks, each containing up to 24 tumors and also 3 cores from benign prostatic tissue as controls. To obtain representative samples, taking the GS into account we arrayed two cores from the primary Gleason pattern and one from the secondary. After morphological control, an additional block was constructed with cores from 5 tumors that were poorly represented in the original blocks. Initially, 333 consecutive tumors were arrayed. However, complete clinical follow-up data could only be collected from 290 men. After exclusion of one case with no cancer in the cores (0.3%) 289 cases remained for analysis, corresponding to 3 x 289 = 867 TMA cores.

The pathogenesis TMA contained seven cores from each case: two from cancer, two from HGPIN and one each from atrophy, non-atrophic benign tissue (without morphological signs of hyperplasia) and benign prostatic hyperplasia (BPH). The BPH sample was taken from TZ and the others from PZ. In the prognostic TMA, 95.3% (826 of 867) of cores contained sufficient amount of cancer for evaluation.

The metastasis TMA included two cores per case. Three cases contained no representative material in the immunostained sections, leaving 17 cases for analysis.
3.2.3 Immunohistochemistry

Immunostaining of TMA sections was performed using commercially available antibodies against PDX-1 (paper V) and HSPs (paper VI). TMA sections were cut at 4 µm, deparaffinized in xylen and rehydrated through graded ethanol. The sections were microwave treated for antigen retrieval (Vector H 3300) for 15 min. Non-specific binding sites were blocked with 5% skimmed milk in PBS with 0.1% BSA. The slides were incubated with the primary antibody at 4°C overnight. Antibodies, dilutions, pretreatment conditions and manufacturers are listed in Table 3. The secondary antibody (Elite kit Vector 5ml/ml in PBS with 0.1% BSA) was applied for 30 min at room temperature. The signal was increased with the ABC kit (20 ml A + 20 ml B/ml PBS-0.1% BSA) at 37°C for 45 min and detected with DAB (Vector kit SK 4100). The slides were counterstained with hematoxylin. Human breast cancer was used as positive and negative control according to instructions from the manufacturer.

<table>
<thead>
<tr>
<th>Antibody/clone</th>
<th>Species/type</th>
<th>Dilution</th>
<th>Retrieval</th>
<th>Vendor</th>
</tr>
</thead>
<tbody>
<tr>
<td>HSP27, 2B4</td>
<td>Mouse/monoclonal</td>
<td>1:20</td>
<td>MW</td>
<td>Novocastra Laboratories</td>
</tr>
<tr>
<td>HSP60, SPA-829</td>
<td>Mouse/monoclonal</td>
<td>1:100</td>
<td>MW</td>
<td>Stressgen Biotechnologies</td>
</tr>
<tr>
<td>HSP70, 8B11</td>
<td>Mouse/monoclonal</td>
<td>1:100</td>
<td>MW</td>
<td>Novocastra Laboratories</td>
</tr>
<tr>
<td>PDX-1</td>
<td>Rabbit/polyclonal</td>
<td>1:4000</td>
<td>MW</td>
<td>Abcam</td>
</tr>
</tbody>
</table>

Table 3: Antibodies/clone, species/type, dilutions, antigen retrieval (MW = microwave treatment) and vendor.

3.2.4 TMA immunostaining evaluation

Intensity and extent of immunoreactivity (IR) in cytoplasm and their product (IRp) were evaluated in each core by two independent observers (S.J. and A.G.) blinded to clinicopathological information. The intensity was scored from 0 (no staining) to 3 (the most intense staining) based on the strongest staining of the core. The extent of positive
cytoplasmatic staining was evaluated semi-quantitatively. Scoring was based on percentage of stained epithelial cells and graded from 0 to 3, signifying 0-4%, 5-25%, 26-50% and >50%, respectively. The mean scores of the two observers were calculated and IRp was defined as the product of these mean values. Thus, an IRp from 0 to 9 was obtained.

3.2.5 Western blot

Seven PCa samples, 3 benign prostatic tissue samples from radical prostatectomy specimens, 2 PCa cell lines (LNCaP, DU145) and a pancreatic cancer cell line (PaCa-2) as positive control were immunoblotted. Cell pellets and the tissue sections were homogenized, total proteins isolated then separated by 1D gel electrophoresis and transferred to a polyvinylidene difluoride membrane. The blot was initially probed for PDX-1 (1:1000, Abcam, Cambridge, MA, USA) then stripped and reprobed for GAPDH (1:1000, Abcam) as protein loading control. Horseradish peroxidase conjugated anti-rabbit (1:3000, GE Healthcare, Piscataway, NJ, USA) secondary antibody was used with GE Healthcare’s enhanced chemiluminescence (ECL) detection kit. Chemiluminescence was detected with a Nikon CCD camera.

3.3 STATISTICAL ANALYSIS (PAPERS I TO VI)

Kappa (k) and Weighted Kappa (wk) are useful measures of interobserver agreement, as the level of agreement is adjusted for that expected by chance. When the observed agreement exceeds chance agreement, k is positive, with its magnitude reflecting the strength of agreement. Thus, k 0.00–0.20 reflects slight agreement, 0.21–0.40 fair agreement, 0.41–0.60 moderate agreement, 0.61–0.80 substantial agreement, and 0.81–1.00 almost perfect agreement. In addition to k, wk uses weights to quantify the relative difference between categories. Slight disagreement (i.e. few steps) is not weighted as heavily as more pronounced disagreement (i.e. several steps). (Landis and Koch 1977; Brennan and Silman 1992) (paper I to VI)

Mann-Whitney U test was used for comparison of non-parametric data. (paper III)

Chi-square test for comparison of proportions. (paper III)
Linear regression according to Pearson was used for univariate regression analysis and a linear multivariate regression model for analysis of multiple explanatory variables. (paper III)

Univariate logistic regression models were used to predict pattern 4/5 carcinoma on prostatectomy specimens. (paper IV)

Spearman's rank correlation was used to determine correlation between biopsy and prostatectomy grade data (paper IV) and correlation between immunoreactivity and GSs and patterns. (paper V and VI)

Differences in means were analyzed using paired and unpaired t-test when appropriate. ANOVA test was used for overall comparison of immunoreactivity and diagnostic categories.

Prognostic parameters were compared by Cox proportional hazards models. P-values less than 0.05 were considered as significant.
4 RESULTS AND DISCUSSION

4.1 REPRODUCIBILITY AND PREDICTION STUDIES (PAPERS I TO IV)

4.1.1 Paper I and III: Interobserver reproducibility of %GG4/5 and mGS in prostatectomy specimens

The four observers reached a overall mean wk for GS, %GG4/5 and mGS of 0.56 (range 0.52-0.66), 0.66 (range 0.58-0.72) and 0.58 (range 0.49-0.74), respectively (table 4). Exact agreement in GS was found in 57% of the cases (range 44% to 70%). In 27% (113/414) of the pairs, there was exact agreement for %GG4/5. In 218 (53%) of the 414 pairs, agreement for mGS was exact.

For %GG4/5 all responses of the four uropathologists were compared pairwise and the differences were calculated. In 40% (164/414), the disagreement was 5% or less, and in 61% (252/414), it was 15% or less. In 8% (35/414), the disagreement was more than 50%. The best agreement for %GG4/5 was between two pathologists working at the same department (wk 0.86). One pathologist (L.E.) reached an intraobserver reproducibility of wk 0.91 for both GS and %GG4/5. In the tumors with greatest disagreement of %GG4/5, crush artifacts, cribriform cancer and high-grade PIN within the tumor were significantly more common. The concordance of %GG4/5 was higher for TZ tumors than for PZ tumors (wk 0.74 and 0.64, respectively; table 5).

<table>
<thead>
<tr>
<th></th>
<th>Average</th>
<th>wk Value (observer no.)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>1</td>
</tr>
<tr>
<td><strong>Prostatectomies</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>GS</td>
<td>0.655</td>
<td>0.558</td>
</tr>
<tr>
<td>%GG4/5</td>
<td>0.720</td>
<td>0.696</td>
</tr>
<tr>
<td>mGS</td>
<td>0.635</td>
<td>0.612</td>
</tr>
<tr>
<td><strong>Prostate biopsies</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>GS</td>
<td>0.552</td>
<td>0.523</td>
</tr>
<tr>
<td>%GG4/5</td>
<td>0.655</td>
<td>0.681</td>
</tr>
</tbody>
</table>

**Table 4**: Weighted kappa values for Gleason scores, percent Gleason grade 4/5 and modified Gleason scores among the four observers.

GS and mGS differed in score distribution (p <0.001, Figure 4). In 183 (66%) of the responses GS and mGS were identical. The difference between GS and mGS was 0, 1
and 2 score units in 66%, 26%, and 8%, respectively, mean 0.41 score units (range 0.24-0.51). Disagreement between observers greater than 1 score unit was more common with mGS than GS (18% and 4%, respectively, p <0.001; figure 5). When

the increment from GS to mGS was 2 units, a tertiary pattern 5 was more often responsible for the score shift than a tertiary pattern 4 (76% and 24%, respectively), while tertiary patterns 5 were less commonly seen in cases with 1 unit increment from GS to mGS (29% and 71%, respectively, p <0.001). Increment was greater for TZ tumors than for PZ tumors (0.63 and 0.35 score units, respectively, p = 0.002). The concordance of mGS was higher for PZ tumors than for TZ tumors (wk 0.59 and 0.52, respectively; table 5). An odd mGS (5, 7 or 9) was more often given than an odd GS (77% and 62%, respectively, p <0.001).

<table>
<thead>
<tr>
<th>wk Value (totals of all observers)</th>
<th>GS</th>
<th>%GG4/5</th>
<th>mGS</th>
</tr>
</thead>
<tbody>
<tr>
<td>Transition zone</td>
<td>0.43</td>
<td>0.74</td>
<td>0.52</td>
</tr>
<tr>
<td>Peripheral zone</td>
<td>0.54</td>
<td>0.64</td>
<td>0.59</td>
</tr>
</tbody>
</table>

Table 5: Weighted kappa values for Gleason score, percent Gleason grade 4/5 and modified Gleason score in different anatomical zones of the prostate.

4.1.2 Paper II: Interobserver reproducibility of %GG4/5 in prostate biopsies

The four observers had a mean wk for GS and %GG4/5 of 0.48 to 0.55 (overall mean 0.51) and 0.52 to 0.68 (overall mean 0.60), respectively (table 4). Exact agreement for
GS was found in 33% to 72% of the comparisons (mean 49%). For %GG4/5 exact agreement was found in 20% to 45% of the comparisons (mean 34%), disagreement of 5% or less in 30% to 61% (mean 43%), disagreement of 15% or less in 42% to 74% (mean 56%) and disagreement of more than 50% in 1% to 39% (mean 20%). There was less disagreement of %GG4/5 when a single biopsy was positive for cancer than in cases where 6 or more biopsies were positive. Number of positive biopsies showed a stronger correlation with discrepancy of %GG4/5 than cancer length. Disagreement was worse when cribriform or fusion patterns were present.

4.1.3 Paper IV: Prediction of %GG4/5

%GG4/5 in prostatectomy specimens was predicted correctly in 34% and within 10%, 20%, and 30% in 55%, 64%, and 73%, respectively. In 53% of the biopsies the %GG4/5 was underestimated and over grading was found in 12% of the cases. Biopsies were classified for %GG4/5 with a sensitivity, specificity and accuracy of 62%, 87% and 69%, respectively. Positive and negative predictive values were 93% and 45%, respectively. Weighted kappa for agreement was slightly higher for GS (0.685) than for %GG4/5 (0.573). The univariate correlation for %GG4/5 in biopsies and the main tumor was $r = 0.62, r^2 = 0.39$ ($p < 0.001$). For %GG4/5, the agreement was better for TZ tumors than in PZ tumors (wk 0.610 and 0.569, respectively) while the opposite was true for the GS (0.617 and 0.663, respectively). %GG4/5 in needle biopsies correlated well with the main tumor in the prostatectomy specimens ($r = 0.62; p < 0.001$). The strongest correlation was obtained for cases with a single biopsy positive for cancer ($r = 0.76$), tumors with a volume less than 2 ml ($r = 0.72$) and for TZ tumors ($r = 0.69$). In univariate logistic regression, %GG4/5 on biopsies predicted the presence of any Gleason grade 4/5 cancer in the main tumor ($p = 0.009$).

4.1.4 Discussion papers I to IV

The Gleason system is the most powerful and most clinically used grading system to date, and predicts prognosis and patient outcome in PCa (Eble JN 2004; Amin, Boccon-Gibod et al. 2005). On the other hand there is a strong tendency toward tumor clustering in the mid range of the GSs with a majority of cases assigned a GS 6 or 7 (Egevad, Granfors et al. 2002; Egevad, Granfors et al. 2002). A recent trend of grade and stage shift aggravates this problem even more (Albertsen, Hanley et al. 2005). In addition tertiary patterns of higher grade proved to have an adverse impact on
biological behavior but the GS provides no solution on how to handle these patterns (Pan, Potter et al. 2000; Stamey, Yemoto et al. 2000; Egevad, Granfors et al. 2002; Hattab, Koch et al. 2006; Patel, Chen et al. 2007).

Some have argued that the Gleason system should be abandoned or modified (Stamey, McNeal et al. 1999; Pan, Potter et al. 2000). Alternative grading methods for PCa have been suggested, namely %GG4/5 and mGS, which both have proven to be independent prognosticators in PCa (Stamey, McNeal et al. 1999; Pan, Potter et al. 2000; Egevad, Granfors et al. 2002; Cheng, Davidson et al. 2007; Vis, Roemeling et al. 2007).

A new grading system should provide prognostic information, which is independent from that given by the GS, reproducibility on prostatectomy and biopsy specimens has to be satisfactory and the new grading must be clinically useful.

In the studies included in this thesis we focused on the interobserver reproducibility (IOR) of the GS and %GG4/5 in prostate biopsies (paper III) and prostatectomy specimens (paper I) and mGS in prostatectomy specimens (paper II). All studies were evaluated by the same four uropathologists. To our knowledge IOR for %GG4/5 and mGS were investigated for the first time in these studies. Difficulties to assess a score with these grading modalities and their influence on interobserver variability are discussed. The ability of needle biopsies to predict prostatectomy %GG4/5 is important for the clinical utility and was analyzed in paper IV.

We were able to demonstrate that IOR of %GG4/5 and mGS were at least as good as that of the GS in the investigated specimen types. We found that wk values of %GG4/5 in prostatectomy specimens were similar to those of needle biopsies (wk 0.66 and 0.60, respectively). However results from the biopsies and prostatectomy specimens must be compared with some caution, as the two series did not include the same patients. Thus, their grade distribution is not necessarily similar.

We compared our results from three grading systems, i.e. GS, %GG4/5 and mGS, using identical prostatectomy series assessed by the same observers. According to wk analysis IOR of mGS in prostatectomy specimens was at least as good as that of the GS, (wk 0.58 and 0.56, respectively) but slightly lower than wk of %GG4/5 (wk 0.66).

A significantly different distribution and a broader variation in grade differences between observers were found with mGS than with GS, possibly because tertiary patterns are not sufficiently well defined. The minimum amount of a pattern that is required for being identified at all is not clear. Occasional incomplete, fused or
branching glands may be interpreted either as cutting artifacts or as a focal pattern 4. Similarly, occasional small solid epithelial structures may be interpreted either as cutting artifacts or as a focal pattern 5. It remains unclear if pattern 5 requires clusters of individual cells, solid strands or solid nests seen at lower than 40X magnification or if this pattern also may be diagnosed when such structures are found interspersed among cancerous glands. Thus, to improve reproducibility of mGS the definition of focal high-grade components has to be clarified. In addition mGS showed a significant shift to odd scores that per se contributed to a clustering in certain scores and thus diminished its prognostic power.

The evaluation of %GG4/5 includes both identification of high-grade tumor and estimation of percent of tumor area involved with such patterns, which may be more difficult to assess than conventional GS or mGS grading. The present thesis demonstrated that the reproducibility of %GG4/5 was at least as good as that of the GS. There are several possible explanations of this finding. When assessing GS and mGS there may be disagreements between GS 5 and 6 and between GS 8 and 9, difficulties possible to avoid by using the %GG4/5. In tumors with homogeneous morphology, a one-step misinterpretation of a Gleason grade results in a two-step error of the GS or mGS as Gleason grade is doubled in such cases. In our prostatectomy series the tumors were distributed among a few GSs or mGSs in the mid range, while all 12 tiers of the %GG4/5 were used. Therefore, a one step error of GS corresponded to a high percentage of the range of GS or mGS, whereas a one step error of %GG4/5 had a smaller influence on the wk. The reproducibility of %GG4/5 in prostatectomy specimens was not improved by dichotomous categorization in tumors with or without Gleason grade 4 to 5. Hence, a problem in reproducing %GG4/5 seemed to be the observer variability of the identification of Gleason grade 4 to 5 patterns rather than the estimation of the percentage of the tumor area occupied by these patterns.

The Gleason grading is based on architectural patterns and it is claimed that high-grade patterns would be more difficult to identify if the amount of tumor is small, e.g. in biopsy specimens. There may be several explanations for the good reproducibility of grading observed in prostate biopsies. They contain a smaller amount of cancer than prostatectomy specimens and this may actually decrease interobserver variability. In our series men with cancer in a single biopsy had more often full concordance of %GG4/5 between the observers (74.4%) than men with at least six biopsies positive for
cancer (18.9%, $p<0.001$). There were also fewer positive biopsies (mean 1.8 and 5.1, respectively, $p=0.002$) and shorter total cancer length (10.1 and 33.7 mm, respectively, $p=0.004$) in cases with full concordance than in the cases with the greatest discrepancy. A possible explanation is that the vast majority of cancers found in single biopsies were assigned a GS 6 similar to small peripheral zone tumors in prostatectomy specimens (McNeal, Villers et al. 1990). Evidently, estimating the proportion of high-grade patterns is easier if many of the cases lack such a component. The greater interobserver variability when judging multiple positive biopsies per case may also reflect that compiling a greater amount of morphological information is more difficult. The disagreement of $\%GG4/5$ shows a stronger correlation with number of positive biopsies than with total cancer length. This suggests that compiling grade data is most difficult when cancer is present in several slides. This source of interobserver variability may be avoided if the last ISUP recommendation is followed - assigning a separate score for each biopsy instead for compiling information in a general GS obtained out of several biopsies. Furthermore, cancer in multiple biopsies may be derived from separate tumor foci with different grades. Also, in larger tumors, there is a greater grade heterogeneity, increasing the likelihood of finding different Gleason patterns in different needle biopsies (Aihara, Wheeler et al. 1994). As in needle biopsies cancer appears in a linear distribution, this may also facilitate the estimation of proportions of grade components. To our knowledge, this is the first study that has compared reproducibility of PC grading with number of positive biopsies and amount of PC in biopsies.

Grading GS and mGS was more difficult in TZ than PZ tumors, while the opposite was true for $\%GG4/5$ grading. In tumors with three or more Gleason grades, the estimation of the proportion of the areas of the grades is decisive for the GS. This is particularly common in TZ tumors with Gleason grades 2, 3, and 4 present in the same tumor resulting in a GS 5, 6 or 7 depending on the amount of each grade. That may explain the lower wk for GS of TZ tumors as compared to PZ tumors. With mGS, this source of interobserver variability is to some extent avoided because a tertiary pattern is included in the score regardless of its proportion of the tumor. This may explain the somewhat better wk results in mGS compared to GS. The converse findings in $\%GG4/5$ may be explained by a significant lower rate of cribriform cancer and high-grade PIN in tumors originating in the TZ. Cribriform cancer is found both in Gleason grades 3 and 4. The distinction between these patterns of cribriform cancer is quite arbitrary and a cause of disagreement among genitourinary pathologists. PIN is not
graded according to the Gleason system, but sometimes it is difficult to distinguish cribriform PIN from invasive cancer. Because of the disagreement about grading of cribriform cancer, the presence of high-grade PIN within the tumor may, therefore, increase the interobserver variability of %GG4/5. Crush artifacts are a potential cause of overgrading of PC because glandular units become disrupted and discohesive epithelial cells appear as solid strands. Crush artifacts are more common close to the surgical margins of prostatectomy specimens and we would expect them to be found more commonly in PZ tumors contributing to the difficulties to assess %GG4/5. However, crush artifacts were only non-significantly more common in PZ tumors in the present series. To our knowledge, this is the first study that has compared the reproducibility of the GS and %GG4/5 between tumors of different zonal origin.

Prostate biopsies are the common standard to guide clinical decision-making. In order to obtain reliable prognostic information before treatment tumor grade must be predictable from preoperative biopsies. We evaluated the correlation of high-grade cancer in a consecutive series of core biopsies and prostatectomy specimens and we showed that sensitivity (62%), specificity (87%) and positive (93%) and negative (45%) predictive values of %GG4/5 on core biopsies were reasonably good if multiple biopsies were taken. In addition there was a correlation between this grading modality in biopsies and prostatectomy specimens for the total amount of cancer (r=0.65, r²=0.42, p<0.001) and the main tumor (r=0.62, r²=0.39, p<0.001).

In the literature we found conflicting results on this subject. Stamey et al. described a strong correlation of %GG4/5 in biopsies and prostatectomy specimens in 120 men (r² = 0.63) (Stamey 1995) and r² = 0.57 in a subset of 89 men (Stamey, McNeal et al. 1999). Rubin et al. studied 101 men with matched biopsies and prostatectomy specimens and showed a rather low sensitivity (38%) but a high specificity (96%) for predicting any pattern 4/5 carcinoma in prostatectomy specimens (Rubin, Mucci et al. 2001). They evaluated the ability of biopsies to predict either any amount of high-grade cancer or ≥10% high-grade cancer in prostatectomy specimens and found a rather weak correlation of %GG4/5 in biopsies vs. prostatectomy specimens (r² = 0.32, p <0.001).

Similar to Rubin et al., we tested the ability to predict both the presence of any pattern 4/5 and also the presence of more than 5% of high-grade cancer. Compared to Rubin et al, we found a stronger correlation for high-grade cancer in biopsies and prostatectomy specimens and also a far higher sensitivity for patterns 4/5.
Different biopsy protocols might explain some of the differences in the results mentioned above. Stamey et al. studied sextant biopsies taken according to Hodge et al., a biopsy strategy that is now widely considered obsolete (Hodge, McNeal et al. 1989). No information was provided by Rubin et al. about number of biopsies taken in each man or the type of biopsy protocol used (although sextant biopsies were mentioned in the article when describing the study strategy, indicating that this might have been the standard protocol at that time). In our study biopsies were taken according to an extended protocol with at least eight biopsies per patient and the mean number of biopsy cores was 10.

The method used for measuring high-grade cancer in biopsies also differed. Rubin et al measured area while we analyzed tumor length, a measure that we believe is more likely to be used in practice. Another difference between the present study and the one by Rubin et al. was that, similar to Stamey et al., we correlated our biopsy findings to the main tumor focus. Current majority opinion is that Gleason grading of PC should be based on the main tumor (Epstein, Allsbrook et al. 2005). When we analyzed %GG4/5 against the total amount of cancer in the prostate, the correlation was slightly stronger than if analyzed against the main tumor.

Conclusions

Before introducing either %GG4/5 or mGS as new grading system their clinical usefulness needs to be proven. Most pathologists are now familiar with the Gleason grading system (Egevad, Allsbrook et al. 2006). Both %GG4/5 and mGS are based on Gleason grading and this could facilitate their introduction in clinical work. %GG4/5 and mGS are based on a simple and all over the world accepted technique, the hematoxylin-eosin staining which is cheap in comparison to other newer techniques. This makes it possible to introduce these new grading systems in a large international scale. It was claimed that the estimation of %GG4/5 or mGS would be more time-consuming than conventional Gleason scoring. However, it probably does not add more to the effort of reporting PC compared to analyses of other novel prognostic markers.

Among arguments against the routine use of %GG4/5 is the fear that this measure would have a poor reproducibility. We demonstrated that the IOR of %GG4/5 on both, needle biopsies and prostatectomy specimens, is at least as high as that of GS (Paper I,
II) and that its sensitivity, specificity, and predictive values on core biopsies are reasonably good if multiple biopsies are taken. There is also a correlation between this grading modality in biopsies and prostatectomy specimens (Paper IV). Others proved the prognostic value of %GG4/5 in PCa (Egevad, Granfors et al. 2002; Cheng, Davidson et al. 2007; Vis, Roemeling et al. 2007). A disadvantage of %GG4/5 is that it does not discriminate GS 6 tumors from GS 4 to 5 tumors, which have a better prognosis according to some studies (Egevad, Granfors et al. 2002). Our opinion is that %GG4/5 remains a theoretically interesting measure of tumor aggressiveness. Whether %GG4/5 will be used in clinical practice remains to be seen.

From studies performed so far, it seems that the IOR of the Gleason grading remains essentially the same with modified Gleason grading and results are probably more influenced by the study design. As an advantage it is shown that mGS minimizes undergrading of prostatic carcinomas in biopsies and improves the agreement between biopsies and radical prostatectomy specimens (Helpap and Egevad 2006). We show in paper III that IOR of mGS is at least as good as that given by the GS. A disadvantage is that mGS has a significantly different distribution and a broader variation in grade differences among observers compared to the GS. In addition mGS leads to a significant shift towards odd scores.

Our study was published in 2004 before the ISUP recommendations in 2005 and might have contributed to the implementation of a modified Gleason grading. But we still think that there are different issues to address despite the new recommendations. For several reasons we believe that it is too early to abandon the conventional GS. Redefining GS as the sum of primary and tertiary patterns without assigning a new term to it is potentially confusing and comparison with results from previous studies using the old definition of GS will turn out difficult. It is of benefit that a new term is introduced to designate the sum of primary and tertiary patterns, tentatively modified GS (mGS), a term that was also used by Pan et al. (Pan, Potter et al. 2000).

The limited number of studies that addressed the prognostic value of a modified GS (Stamey, McNeal et al. 1999; Pan, Potter et al. 2000; Stamey, Yemoto et al. 2000; Egevad, Granfors et al. 2002; Helpap and Egevad 2006; Veloso, Lima et al. 2007; Helpap in press; Helpap in press) is insufficient to support the introduction of a new grading system. Before recommending a general use of mGS, previous results on its prognostic value need to be corroborated.
We show that the IOR of mGS on prostatectomy specimens was at least as high as that of GS (Paper III) but its reproducibility on prostate biopsies should be further evaluated. Additionally, reproducibility of mGS has to be improved and consensus must be reached on the definition of tertiary Gleason patterns. Furthermore, the ability of preoperative core biopsies to predict mGS in prostatectomy specimens is described only in a few studies (Helpap and Egevad 2006; Veloso, Lima et al. 2007) and needs to be further investigated. Theoretically, a minimal component of tertiary pattern would easily go undetected by needle biopsies, decreasing the practical utility of a mGS.

In conclusion, there is a gap between the relatively small amount of published data on mGS and its role as a newly introduced prognosticator in PCa and there is an urgent need of further studies.

4.2 PROTEIN EXPRESSION STUDIES (PAPERS V AND VI)

4.2.1 Paper V: Pancreatic duodenal homeobox-1

PDX-1 expression was characterized in benign and malignant prostatic tissues. For this we analyzed TMAs for studies of prognosis, pathogenesis and metastases as described in chapter 3.2.2.

PDX-1 was expressed in cytoplasm and occasional nuclei of epithelial cells of benign non-atrophic glands, PCa, HGPIN and atrophic glands (figures 9 to 12, page 54).

Distributions of PDX-1 immunoexpression in prostatic tissue and lymph node metastases were calculated as IRp (IRp = the product of immunoreactivity intensity and extent) as follows:

![Figure 6: PDX-1 expression in benign prostatic tissue, atrophy, PIN and primary and secondary PCa.](image-url)
PDX-1 was overexpressed not only in cancer vs. benign tissue but also in atrophy and HGPIN compared to cancer and benign non-atrophic prostatic epithelium. The expression was stronger in benign non-atrophic glands of PZ than in hyperplastic glands of TZ. There was no overexpression in lymph node metastases compared to primary carcinomas.

<table>
<thead>
<tr>
<th></th>
<th>Benign non-atrophic</th>
<th>Benign PZ</th>
<th>Benign TZ</th>
<th>PCa</th>
<th>HGPIN</th>
<th>Atrophy</th>
</tr>
</thead>
<tbody>
<tr>
<td>Benign non-atrophic</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>&lt;0.001</td>
<td>&lt;0.001</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>Benign PZ</td>
<td>-</td>
<td>&lt;0.001</td>
<td>&lt;0.001</td>
<td>&lt;0.001</td>
<td>&lt;0.001</td>
<td></td>
</tr>
<tr>
<td>Benign TZ</td>
<td>-</td>
<td></td>
<td>&lt;0.001</td>
<td>&lt;0.001</td>
<td>&lt;0.001</td>
<td></td>
</tr>
<tr>
<td>PCa</td>
<td>-</td>
<td></td>
<td></td>
<td>0.022</td>
<td>&lt;0.001</td>
<td></td>
</tr>
<tr>
<td>HGPIN</td>
<td>-</td>
<td></td>
<td></td>
<td></td>
<td>0.102</td>
<td></td>
</tr>
<tr>
<td>Atrophy</td>
<td>-</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Table 6: Correlations of PDX-1 immunoexpression in prostatic tissues (p values, paired t-test).

To validate our results, specifically the differential expression of PDX-1 observed in benign vs. PCa tissue, we analyzed human tissue samples and carcinoma cell lines by Western blot. The presence of PDX-1 in human prostatic tissue was confirmed. The benign tissue samples displayed a weaker signal in comparison to cancer samples and PCa cell lines.

To our knowledge, only Wang et al. (Wang, Li et al. 2005) reported on PDX-1 expression in PCa. Using TMA, they also found a stronger PDX-1 expression in PCa than in benign prostatic tissue. However, our study also included HGPIN and prostatic atrophy, showing an even stronger PDX-1 expression. In some other types of cancer PDX-1 was proposed to be an important regulator of tumor cell proliferation and differentiation (Koizumi, Doi et al. 2003; Wang, Li et al. 2005). In view of this, the observation that PDX-1 is overexpressed not only in cancer but also at even higher level in prostatic atrophy and HGPIN is interesting. HGPIN is an established precursor of invasive cancer and it is postulated that atrophy may play a role in carcinogenesis (De Marzo, Platz et al. 2007).
Thus, PDX-1 overexpression in these lesions may indicate the involvement of this protein in early carcinogenesis of the prostate.

In the prognostic TMA, 289 consecutive radical prostatectomy specimens were analyzed for PDX-1 staining (IRp). The PDX-1 expression did not correlate with either biochemical recurrence or GS (p = 0.19 and 0.37, respectively).

The expression levels in cores of primary and secondary Gleason patterns were analyzed separately against outcome data, but did not correlate either (data not shown). Thus, despite the up regulation of PDX-1 in PCa and its precursors, no correlation with prognosis was found.

The results suggest that PDX-1 may play a role in the pathogenetic development of PCa but lacks significance as predictor of progression of invasive carcinoma. Future studies will further clarify the role of PDX-1 in neoplastic transformation in the prostate.

4.2.2 Paper VI: Heat shock proteins 27, 60 and 70

HSP27, HSP60 and HSP70 expression were evaluated in benign and malignant prostatic tissues. We examined these proteins with respect to clinicopathological factors and prognosis using the same prognostic TMA constructed from 289 consecutive radical prostatectomy specimens as described in the PDX-1 study.

HSPs were expressed in cytoplasm and occasionally in nuclei of secretory cells in benign and malignant prostate glands. Benign prostate tissue sometimes showed staining of basal cells. The cytoplasmic staining pattern varied from diffuse non-granular to fine granular and in some cases it showed a membranous pattern. The cytoplasmic staining for HSP60 was more coarse granular than that of the other HSPs (Figures 13 and 14, page 54).

Distributions of HSPs immunoexpression were calculated as IRp and mean IRp for each of the proteins depicted in Figure 7. Strong expression (IRp ≥6) was seen in 31 (10.7%), 36 (12.5%), and 110 (38.1%) of cases when stained against HSP27, HSP60 and HSP70, respectively. For HSP27 and HSP60, but not HSP70, increasing IRp correlated significantly with increasing Gleason score.
The expression (IRp) of HSP27 and HSP60, but not of HSP70 significantly anticipated biochemical recurrence in univariate Cox analysis (p = 0.014, 0.034, and 0.160, respectively).

Recurrence-free survival in patients with strong expression of HSP27 and HSP60 (IRp ≥6) was shorter than in those with weak expression (p = 0.019 and 0.001, respectively).

In multivariate analysis, only HSP60 IRp remained an independent predictor of biochemical recurrence when extraprostatic extension, positive surgical margins, seminal vesicle invasion, and Gleason score were included in the model as explanatory variables.
<table>
<thead>
<tr>
<th>Prognostic factor</th>
<th>HR (CI)</th>
<th>p value</th>
</tr>
</thead>
<tbody>
<tr>
<td>HSP60</td>
<td>2.28 (1.40-3.71)</td>
<td>0.001</td>
</tr>
<tr>
<td>Gleason score</td>
<td>1.36 (1.08-1.72)</td>
<td>0.010</td>
</tr>
<tr>
<td>Extra-prostatic extension</td>
<td>1.13 (0.69-1.84)</td>
<td>0.63</td>
</tr>
<tr>
<td>Margin status</td>
<td>2.91 (1.81-4.68)</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>Seminal vesicle invasion</td>
<td>1.88 (1.04-3.40)</td>
<td>0.036</td>
</tr>
</tbody>
</table>

**Table 7:** Multivariate Cox analysis of HSP60 IRp (the product of immunoreactivity intensity and extent) ≥6 vs. <6. Hazard ratio (HR) with 95% confidence interval (CI).

However, if preoperative clinical data (preoperative serum PSA and age at surgery) were added to the model, only surgical margins (p <0.001), Gleason score (p = 0.004), and serum PSA (p <0.001) remained significant.

The published data on HSP expression in cancer are complex. Cancers may show either over- or underexpression of HSPs compared with benign tissue. Furthermore, the prognostic impact of HSP expression varies between tumor types and correlates with either prolonged or shortened survival.

For **HSP27**, Thomas et al. described a variable expression with apparent loss of staining intensity as the Gleason scoring increased and the tumor became more invasive (Thomas, Brown et al. 1996). However, a small set of frozen prostate samples was used including 13 cancers and 15 benign cases. In contrast, we showed that strong HSP27 expression correlated with biochemical recurrence and Gleason score. Our results were in line with two other studies on PCa (Cornford, Dodson et al. 2000; Kurahashi, Miyake et al. 2007). Cornford et al. found in univariate analysis of 85 advanced cancers (T3-4, Nx, M0-1) a correlation between high HSP27 expression and poor clinical outcome. They also reported a decreased HSP27 expression in 25 early PCa (T1-2NoMo) (Cornford, Dodson et al. 2000). Kurahashi et al. found an association between high HSP27 expression in 172 cases of PCa and pathological stage, Gleason score, surgical margin status, lymph node metastasis, tumor volume, and a shorter biochemical recurrence-free survival (Kurahashi, Miyake et al. 2007). Our results confirm their findings that HSP27 was not an independent predictor of biochemical recurrence in multivariate analysis and that recurrence-free survival in patients with
strong HSP27 expression was significantly lower than that in those with weak expression. 

For **HSP60** Cornford et al. showed an elevated expression in both HGPIN and PCa compared to benign tissue. They did not find any correlation with Gleason grade or clinical outcome, but these analyses were based on a smaller series of cases than ours. Johansson et al. found an upregulation of HSP60 in PCa cell lines and in tumor tissue from 79 men. However, without doing statistical analysis they noticed no apparent relation between Gleason score and intensity (Johansson, Pourian et al. 2006). Our data indicated that HSP60 overexpression correlated with both biochemical recurrence and Gleason score. In multivariate analysis, we found for the first time HSP60 to be an independent predictor of biochemical recurrence when EPE, positive surgical margins, SVI, and Gleason score were included in the model.

Abe et al. compared **HSP70** plasma levels in patients with untreated hormone-refractory PCa and healthy controls. Overall, plasma HSP70 did not predict PCa diagnosis more accurately than serum PSA, but raised levels were seen in several PCa patients with normal serum PSA (Abe, Manola et al. 2004). HSP70 stabilizes mutated p53, indicating that HSP70 may have relevance for PCa progression (Hinds, Finlay et al. 1987). However, in the current study, we saw no correlation between HSP70 and biochemical recurrence or with histopathological prognostic factors. These results are in line with other studies, which showed no correlation with Gleason grade or clinical outcome (Cornford, Dodson et al. 2000; Kurahashi, Miyake et al. 2007).

Our research concentrated on the prognostic value of HSPs but of equal importance is their clinical relevance as potential therapeutic targets in PCa patients. It was described that inhibiting the function of HSPs in tumor cells might be an attractive strategy in cancer treatment (Lebret, Watson et al. 2003; Sharp and Workman 2006). Thus, it is clear that monitoring the expression level of HSPs in human PCa may be of clinical interest in the future.

4.2.3 Reproducibility of TMA staining evaluation

Immunohistochemistry on TMA is a now commonly used method, but the optimal assessment routine needs to be further evaluated. To our knowledge, no systematic studies have been done on the interobserver reproducibility of the interpretation of immunohistochemical stains of PCa TMAs.
We calculated the wk values for interobserver reproducibility of immunoreactivity intensity and extent and IRp (the product of intensity and extent) for both HSPs and PDX-1. The results are summarized in Table 8.

<table>
<thead>
<tr>
<th>Marker</th>
<th>Intensity</th>
<th>Extent</th>
<th>IRp</th>
</tr>
</thead>
<tbody>
<tr>
<td>HSP27</td>
<td>0.823</td>
<td>0.244</td>
<td>0.719</td>
</tr>
<tr>
<td>HSP60</td>
<td>0.713</td>
<td>0.036</td>
<td>0.641</td>
</tr>
<tr>
<td>HSP70</td>
<td>0.613</td>
<td>0.077</td>
<td>0.584</td>
</tr>
<tr>
<td>PDX-1</td>
<td>0.650</td>
<td>0.130</td>
<td>0.540</td>
</tr>
</tbody>
</table>

Table 8: Weighted kappa values (wk) for interobserver reproducibility of immunoreactivity intensity, extent and IRp (the product of intensity and extent).

Interobserver reproducibility of the staining intensity varied between being substantial and almost perfect. Estimating both intensity and extent is a common method for assessment of immunostains. However, estimation of extent presents several difficulties. The distribution of immunoexpression of proteins is typically not strictly dichotomous with cells being either positive or negative. Cancer often displays a continuous range of staining intensities making it difficult to estimate exact proportions. Another problem is that it is unclear whether focally strong staining or a diffuse staining of lower intensity has the greatest impact on tumor behavior. Furthermore, the necessity of scoring IR extent in TMAs is questionable as this measure is indirectly factored in when intensity is estimated in multiple individual cores from the same tumor.

In the current studies, the main conclusions remained unchanged regardless of whether staining extent was included in the analysis (data not shown).
Figure 9:
Immunohistochemical staining for PDX-1. Weak staining in benign non-atrophic glands. (Original lens magnification 20X)

Figure 10:
Immunohistochemical staining for PDX-1. Strong staining in PCa, Gleason score 3+3=6. (Original lens magnification 20X)

Figure 11:
Immunohistochemical staining for PDX-1. Strong staining in HGPIN. (Original lens magnification 20X)

Figure 12:
Immunohistochemical staining for PDX-1. Strong staining in atrophic glands. (Original lens magnification 20X)

Figure 13:
Immunohistochemical staining for HSP27. Patchy cytoplasmic staining in PCa, Gleason score 3+3=6. Benign prostatic glands (arrow) sometimes showed staining of basal cells. (Original lens magnification 40X)

Figure 14:
Immunohistochemical staining for HSP60. Strong cytoplasmic staining in PCa, Gleason score 3+3=6. The cytoplasmic staining for HSP60 was more coarsely granular than that of the other HSPs and in average more intense. (Original lens magnification 40X)
5 ACKNOWLEDGEMENTS

I would like to express my sincere gratitude to the Karolinska Institute and Karolinska Hospital for giving me the opportunity to perform my research, and to all my colleagues, my family and friends who have supported me throughout the work of this thesis. Especially I would like to thank to:

Associate professor Lars Egevad, a great supervisor, for accepting me as a PhD student in his research group and introducing me in the world of research. I thank you for your outstanding support during all these years. You have always found the time to discuss with me even when your time was limited and the geographical distance was huge. I will miss our frequent e-mail conversations, especially the detailed humorous reports from all over the world.

The members of the Egevad group: Sara Jonmarker, Lars Häggarth and Alexander Valdman for being always helpful.

My colleagues Birgitta Sundelin, Hans Hamberg, Carl-Gustav Pihl. Without your engagement the first three articles wouldn’t have been possible.

All my other co-authors.

Anna Lindberg and Nicole Lyandrat for providing excellent lab work.

My colleagues from the Danderyd and Solna section of the pathology department, for giving me support and the air to breathe when the pressure was high.

My colleague Lars Häggarth for revising my thesis.

My friend Jens Habermann for fruitful discussions and for reviewing my thesis.

My family and especially my parents Helga and Jochen Glaessgen and my brother Christoph Glaessgen for your unconditioned love and support during all these years.

And last but not least my wife Daria Glaessgen and our children Ioana and Felix for standing by my side during this time, being a source of inspiration and joy.
6 REFERENCES


Eble JN, S. G., Epstein JI, Seste...


Montironi, R., C. M. Galluzzi, et al. (1993). "Prostatic intra-epithelial neoplasia: expression and location of proliferating cell nuclear antigen in epithelial,


