1. Introduction

This Summary of Safety and Performance (SSP) is intended to provide public access to an up-to-date summary of the main aspects of the safety and performance of the device.

The SSP is not intended to replace the Instructions For Use as the main document to ensure the safe use of the device, nor is it intended to provide diagnostic or therapeutic suggestions to intended users.

The following information is intended for professional users.

2. Scope

This Summary of the Safety and Performance (SSP) is valid for the following products

(a) Prostatype RT-qPCR kit

- Article number PG0009 / UDI-DI 07350086460094
- Article number PG0010 / UDI-DI 07350086460100

(b) P-score Web Service (PWS)

- Article number PG0004 / UDI-DI 07350086460049

3. Responsibility

The Head of Regulatory Affairs and Quality Assurance is responsible for the upkeep and maintenance of the document.

4. Abbreviations

PWS	P-score Web Service
RT-qPCR	Reverse transcription quantitative polymerase chain reaction
FFPE	Formalin-fixed paraffin-embedded
PCa	Prostate cancer
PSA	Prostate specific antigen
GAPDH	Glyceraldehyde-3-phosphate dehydrogenase
IGFBP3	Insulin-like growth factor binding protein 3
F3	Voagulation factor III
VGLL3	VSestigial-like family member 3
PCSM	Prostate cancer specific mortality
AUC	The Area Under the ROC Curve
OS	Overall survival

CRPC	Castration-resistant Prostate cancer

5. Summary of safety and performance

I. Device identification and general information

I.I Device trade name(s)

(a) Prostatype RT-qPCR kit

(b) P-score Web Service (PWS)

I.II Manufacturer's name and address

Prostatype Genomics AB

Augustendalsvägen 20

131 52 Nacka Strand

SWEDEN

I.III Manufacturer's single registration number (SRN)

SE-MF-000013349

I.IV Basic UDI-DI

(a) The Prostatype® RT-qPCR Kit: 735008646PG000526

(b) P-score Web Service: 735008646PG000424

I.V European Medical Device Nomenclature (EMDN) description / text:

(a) Prostatype kit:

- EMDN level 3 code: W0106 - Genetic testing
- Most granular level: W010699 - genetic testing other

(b) P-score Web Service (PWS):

- EMDN level 3 code: W0201 - Chemistry / Immunochemistry Instruments

- Most granular level: W0201030192 - MIXED PANEL MULTIPARAMETER ANALYSERS - IVD MEDICAL DEVICE SOFTWARE

I.VI Risk class of device

Class C (rule 3h)

I.VII Indication whether it is a device for near-patient testing and/or a companion diagnostic

The device is neither for a near-patient testing nor a companion diagnostic.

I.VIII Year when the first certificate was issued under Regulation (EU) 2017/746 covering the device

2024

I.IX Authorized representative if applicable; name and the SRN

Not Applicable.

I.X NB's name (the NB that will validate the SSP)

TÜV SÜD: NB 0123

II. Intended use of the device

II.I Intended purpose (elements in Annex II 1.1 (c))

(a) Prostatype RT-qPCR kit

Prostatype RT-qPCR kit is an In Vitro Diagnostic (IVD) medical device analyzing the mRNA expression levels of three embryonal cancer stem cell genes in existing biopsies of patients who have been diagnosed with localized or metastasis-free prostate cancer. The kit is to be used in conjunction with P-score Web Service (PWS), to provide additional prognostic information to guide treatment decisions. The P-score results shall be used in conjunction with other clinical tools for prognostic evaluation, and not used as an independent predictor. The Prostatype RT-qPCR kit and the P-score are not intended to be used as diagnosis of prostate cancer.

(b) P-score Web Service (PWS)

P-score Web Service (PWS) is an In Vitro Diagnostic (IVD) medical device software intended to be used to calculate a P-score, based on gene expression data from the Prostatype RT-qPCR kit together with other established clinical data (PSA, Gleason score, & Tumor stage). P-score provide additional prognostic information to guide treatment decisions of patients who have been diagnosed with localized or metastasis-free prostate cancer. The P-score results shall be used in conjunction with other clinical tools for prognostic evaluation, and not used as an independent predictor. The Prostatype RT-qPCR kit and the P-score are not intended to be used as diagnosis of prostate cancer.

II.II Indication(s) and target population(s)

Indications: Localized metastasis free Prostate Cancer

Intended patient population:

(a) Prostatype RT-qPCR kit

Prostatype RT-qPCR kit is intended for the prognostic evaluation by physicians as a decision support tool for evaluating treatment options for patients over the age of 18 with localized prostate cancer.

(b) P-score Web Service (PWS)

P-score Web Service (PWS) is intended for the prognostic evaluation by physicians as a decision support tool for evaluating treatment options in patients over the age of 18 with localized prostate cancer.

II.III Limitations and/or contra-indications (e. g. relevant interferences, cross-reactions)

Limitations: Men over the age of 18 years

Contra-indications: Patients with metastasis at diagnosis

III. Device description

III.I Description of the device, including the conditions to use the device (e.g. laboratory, near-patient testing)

(a) Prostatype® RT-qPCR Kit

Prostatype RT-qPCR kit is an In Vitro Diagnostic (IVD) medical device including a PCR test, which utilizes one-step reverse transcription quantitative PCR (RT-qPCR) for the assessment of the mRNA expression levels of three genes (IGFBP3, F3, and VGLL3) relative to the expression level

of GAPDH in total RNA extracted from formalin-fixed paraffin-embedded (FFPE) human prostate cancer tissue.

(b) P-score web service (PWS)

P-score Web Service (PWS) is a cloud based medical device software that calculates a risk score, P-score, based on the expression levels of three genes (IGFBP3, F3, and VGLL3) relative to the expression level of GAPDH, from the Prostatype RT-qPCR kit, and other established clinical data (PSA, Gleason score, & Tumor stage). The P-score provide prognostic insights to guide treatment decisions of patients diagnosed with localized prostate cancer.

The principal of operation and the mathematical approach

PWS calculates a risk score (P-score) to predict the risk of death from prostate cancer. P-score is a point-based scoring system based on a Fine-Gray competing risk model. The model was used to establish the P-score based on the three-gene signature combined with PSA, Gleason grade, & Tumor stage at diagnosis.

III.II In case the device is a kit, description of the components (including regulatory status of components, for example, IVDs, medical devices and any Basic UDI-DIs)

(a) Prostatype® RT-qPCR Kit

Prostatype RT-qPCR kit consists of various reagents. Only the Prostatype kit is CE marked not the raw materials included in the kit. All the reagents go through incoming quality control check and must meet required criteria to be incorporated in the kit.

(b) P-score web service (PWS)

P-score web service (PWS) is not a kit.

III.III A reference to previous generation(s) or variants if such exists, and a description of the differences

(a) Prostatype® RT-qPCR Kit

Prostatype RT-qPCR kit is Product Model number 1.

(b) P-score web service (PWS)

P-score web service (PWS) is Product Model number 2.

PWS had an older generation called CPMA which was a stationary software that needed to be installed in the computer whereas PWS is a cloud-based solution.

III.IV Description of any accessories which are intended to be used in combination with the device

Devices have no accessories.

III.V Description of any other devices and products which are intended to be used in combination with the device

Need RT-qPCR instrument, FFPE RNA extraction kits and general laboratory accessories are required. Detailed information can be found in the kit's instructions for use.

IV. Reference to any harmonized standards and CS applied

Common specifications	
IEC-62304:2006 + AMD1:2015 *For P-score Web Service (PWS)*	Medical device software – Software life cycle processes + Amendment 1 - Medical device software - Software life cycle processes
Harmonised standards	
EN ISO 13485:2016, **EN ISO 13485:2016/AC:2018,** **EN ISO 13485:2016/A11:2021**	Medical devices - Quality management systems - Requirements for regulatory purposes
EN ISO 14971:2019, **EN ISO 14971:2019/A11:2021**	Medical devices - Application of risk management to medical devices
EN ISO 15223-1:2021	Medical devices - Symbols to be used with information to be supplied by the manufacturer - Part 1: General requirements

V. Risks and warnings

V.I Residual risks and undesirable effects

(a) Prostatype RT-qPCR kit

The overall residual risk posed by Prostatype® RT qPCR Kit has been demonstrated to be acceptable using the criteria defined in RMP-1 Risk Management Plan, Prostatype® RT qPCR Kit

(b) P-score web service (PWS)

The overall residual risk posed by P-score Web Service has been demonstrated to be acceptable using the criteria defined in the RMP-3 Risk Management Plan, P-score Web Service

V.II Warnings and precautions:

(a) Prostatype RT-qPCR kit, included in IFU B0001

To protect yourself as well as to avoid contamination of reagents and samples, always wear a laboratory coat and disposable gloves when using the Prostatype® RT-qPCR kit. This measure is also critical when scraping cancer tissue for RNA extraction.
NOTE: The Prostatype® RT-qPCR Kit does not contain any hazardous ingredients. The product

does not contain any infectious substances or agents, which cause disease in humans or animals.
NOTE: The respective safety data sheets (SDS) are available on our website and provide more detailed information (https://p-score.prostatype.se/documents, available for certified test labs)

To obtain reliable results, it is imperative to adhere to the instructions given in the IFU. Deviations from these guidelines may cause the generation of incorrect or invalid data or may lead to the failure to obtain test results.

General:

● The Prostatype RT-qPCR kit is for use in a professional laboratory and only be used by professional laboratory personnel that has been trained and certified by Prostatype Genomics according to the corresponding training document.
● Avoid microbial contamination of reagents when removing aliquots from reagent bottles.
● To prevent contamination by amplicons generated from previous PCR reactions, a strict separation of pre-PCR activities (e.g., RNA extraction, PCR setup) and PCR activities (e.g. real-time PCR) is strongly recommended. Preferably, a working area exclusively used for RNA testing shall be used.
● General recommendation on the organization and procedures of the laboratory shall be adhered to in order to prevent DNA contamination. Always wipe all working areas, surfaces, equipment and materials with a DNA decontamination agent and an RNase decontamination agent before starting the procedure.

Patient Sample handling and RNA extraction

● Samples intended for analysis with the Prostatype RT- qPCR kit are to be selected by a professional pathologist or a professional with equivalent training.
● An optimal amount of 30 mm2 tissue containing at least 50% cancer cells is recommended for each patient sample.
● Good laboratory practice is essential to minimize the risk of cross-contamination between samples during and after RNA extraction and purification.
● The use of sterile pipette tips with an aerosol barrier is recommended to prevent cross-contamination of patient samples.

PCR activities

● Do not store RNA after extraction and ensure that freshly extracted RNA is used in the Prostatype® RT-qPCR test within one day.
● Preparation of the Prostatype® RT-qPCR shall be performed at room temperature (15-25 °C).
● We further recommend that used PCR plates should be placed in a re-sealable plastic bag or equivalent immediately after removal from the PCR instrument. The bag shall be closed and disposed of in a dedicated waste container. Never store a used PCR plate outside of the PCR instrument. Never open a used PCR plate.

Disposal

Follow local disposal regulations based on your location along with recommendations and content in the Safety Data Sheet to determine the safe disposal of the products.

(b) P-score web service (PWS), included in IFU B0002

WARRANTY
Wrong inclusion of data when calculating P-score may affect the results, in which event Prostatype Genomics AB disclaims all warranties expressed, implied or statutory, including the implied warranty of merchantability and fitness for use. Prostatype Genomics AB and its authorized distributors, in such event, shall not be liable for consequential.

Note: It is advised to log-out of the PWS after use and to keep a good cyber hygiene in general.

V.III Other relevant aspects of safety, including a summary of any field safety corrective action (FSCA including FSN), if applicable

There has not been any field safety corrective action associated with the devices.

VI. Summary of performance evaluation and post-market performance follow-up (PMPF)

Internal scientific validity findings (company sponsored studies) for genomic biomarkers applied in

(a) Prostatype® RT-qPCR kit

(b) P-score web service (PWS)

Study title	*An expression signature at diagnosis to estimate prostate cancer patients' overall survival*[1]
Author	*Zhuochun Peng*
Journal & Publication year	*Prostate Cancer Prostatic Diseases, 2014*
DOI	*10.1038/pcan.2013.57*
AIM	*This study aimed to identify biomarkers for estimating the overall and prostate cancer (PCa)-specific survival in PCa patients at diagnosis.*
Cohort	*Swedish cohort of 189 PCa patients diagnosed between 1986 and 2001*
Method	*To explore the importance of embryonic stem cell (ESC) gene signatures, the authors identified 641 ESC gene predictors (ESCGPs) using published microarray data sets. Then, by using an independent dataset containing 112 PCa tissue samples, they narrowed down the list of candidate genes to 258 ESCGPs. Of these, 34 genes were selected by their high-rank order in subtype classification or by their ability to discriminate between tumor samples and normal tissue samples, as well as with robust performance in RT-qPCR reaction. Selected genes were analyzed by multiplex quantitative polymerase chain reaction using prostate fine-needle aspiration samples. Univariate and multivariate Cox proportional hazard ratios and Kaplan–Meier plots were used for the survival analysis*
Results	*An expression signature of VGLL3, IGFBP3 and F3 was shown sufficient to categorize the patients into high-, intermediate- and low-risk subtypes. The median overall survival times of the subtypes were 3.23, 4.00 and 9.85 years, respectively. The difference corresponded to hazard ratios of 5.86 (95% confidence interval (CI): 2.91–11.78, Po0.001) for the high-risk subtype and 3.45 (95% CI: 1.79–6.66, Po0.001) for the intermediate-risk compared with the low-risk subtype.*

Study title	Improving the Prediction of Prostate Cancer Overall Survival by Supplementing Readily Available Clinical Data with Gene Expression Levels of IGFBP3 and F3 in Formalin-Fixed Paraffin Embedded Core Needle Biopsy Material [2].
Author	*Zhuochun Peng*
Journal & Publication year	PLOS one, 2015
DOI	*10.1371/journal.pone.0145545*
AIM	*Verification the prognostic value of previously identified gene expression signature in formalin fixed paraffin embedded (FFPE) prostate core needle biopsy tissue samples.*
Cohort	*241 prostate cancer patients diagnosed from 2004–2007 with a follow-up exceeding 6 years, a Swedish Cohort.*
Method	*The cohort consisted of four patient groups with different survival times and death causes. A four multiplex one-step RT-qPCR test kit (Prostatype RT-qPCR kit), designed and optimized for measuring the expression signature in FFPE core needle biopsy samples, was used. In archived FFPE biopsy samples the expression differences of two genes (IGFBP3 and F3) were measured. The survival time predictions using the current clinical parameters only, such as age at diagnosis, Gleason score, PSA value and tumor stage, and clinical parameters supplemented with the expression levels of IGFBP3 and F3, were compared.*
Results	*The cohort was composed of two main groups: deceased groups and alive groups, with similar number of patients in each main group. For deceased groups, there were the prostate cancer-specific deceased group (PCa group), and the deceased group who died of other diseases (OD group) with short median survival time 2.6 and 2.1 years respectively. The matched alive group (MA group) and the alive group randomly selected (RA group) were two alive groups with longer median survival time 7.7 and 7.5 years respectively. The MA patients had higher ages (71.0 years) at diagnosis as compared to RA patients (67.1 years) with P value 0.0150; the two groups with deceased patients also had higher ages at the time of diagnosis (PCa, 73.2 years; OD, 76.0 years). ROC curves for grouping variable prediction were estimated to show the sensitivity and the specificity of survival prediction. As compared to the prediction model that used only the clinical parameters, the addition of the expression levels of IGFBP3 and F3 increased the prediction performance in three groups, except in the PCa group with a slightly lower performance. The Area Under the ROC Curve (AUC) value was increased in OD, MA and RA*

	patient groups. Particularly for the two alive patient groups with much longer survival time, AUC value was increased from 0.627 to 0.690 and from 0.787 to 0.848, respectively.

Study title	*A Novel Risk Score (P-score) Based on a Three-Gene Signature, for Estimating the Risk of Prostate Cancer-Specific Mortality*[3]
Author	*Fabian Söderdahl*
Journal & Publication year	*Research and Reports in Urology, 2022*
DOI	*10.2147/RRU.S358169*
AIM	*To develop and validate a risk score (P-score) algorithm which includes previously described three-gene signature and clinicopathological parameters to predict the risk of death from prostate cancer (PCa) in a retrospective cohort.(P-score algorithm is the basis of PWS).*
Cohort	*591 PCa patients diagnosed between 2003 and 2008 in Stockholm, Sweden, with a median clinical follow-up time of 7.6 years*
Method	*Expression of a three-gene signature (IGFBP3, F3, VGLL3) was measured in FFPE material from diagnostic core needle biopsies (CNB) of these patients using Prostatype RT-qPCR kit. A point-based scoring system based on a Fine-Gray competing risk model was used to establish the P-score based on the three-gene signature combined with PSA value, Gleason score and tumor stage at diagnosis. The endpoint was PCa-specific mortality, while other causes of death were treated as a competing risk. Out of the 591 patients, 315 patients (estimation cohort) were selected to develop the P-score. The P-score was subsequently validated in an independent validation cohort of 276 patients.*
Results	*The P-score was established ranging from the integers 0 to 15. Each one-unit increase was associated with a hazard ratio of 1.39 (95% confidence interval: 1.27–1.51, $p < 0.001$). The P-score was validated and performed better in predicting PCa-specific mortality than both D'Amico (0.76 vs 0.70) and NCCN (0.76 vs 0.71) by using the concordance index (C-index) for competing risk. Similar improvement patterns are shown by time-dependent area under the curve. As demonstrated by cumulative incidence function, both P-score and gene signature stratified PCa patients into significantly different risk groups. Additionally, P-score remained significantly associated with an increased risk for PCa-specific mortality in a bivariate analysis ($p < 0.001$), this was not true for D'Amico or NCCN.*

Validation studies for evaluating clinical performance of the biomarker genes in

(a) Prostatype® RT-qPCR kit

(b) P-score web service (PWS)

Study title	*Validation of the prognostic value of a three-gene signature and clinical parameters-based risk score in prostate cancer patients (proposed title can be changed during publication)*
Author	Arni Saemundsson
Journal & Publication year	*Prostate, 2023*
DOI	10.1002/pros.24530.
AIM	*To validate the prognostic value of Prostatype® risk score (P-score) which includes the gene signature and conventional risk factors in a retrospective cohort.*
Cohort	*716 patients diagnosed with PCa from 2008 to 2010 in Skåne University Hospital in Sweden with a median follow-up of 8.8 years.*
Method	*After excluding patients based on pathological and clinical eligibility criteria, RNA quality, and presence of metastasis at diagnosis, a final cohort comprising 316 patients was further analyzed. Expression levels of IGFBP3, F3, and VGLL3 were determined from archived FFPE core needle prostate biopsy tissues. The gene expression data was combined with clinical parameters (Gleason score, PSA, and clinical T-stage) to estimate the P-score. P-score prediction on prostate cancer specific mortality (PCSM), distant metastasis and adverse pathological outcomes were investigated.*
Results	*P-score effectively stratified the patients into three risk groups (low, intermediate, and high). P-score predicted both PCSM (HR=1.6) and metastasis (HR=1.46). P-score had an AUC of 0.93 in predicting the PCSM risk at 10 years (95% CI: 0.89-0.98) significantly better than both ĎAmico (AUC: 0.81, 95% CI:0.72-0.90, p<0.001), and UCSF-CAPRA (AUC: 0.88, 95% CI:0.80-0.96, p<0.05). Decision-curve analysis showed a higher net benefit of P-score compared to both D'Amico and CAPRA. For radical prostatectomy treated patients, higher P-score was greatly associated with adverse pathological features such as pathologic stage T3-4, ISUP grades ≥ 3. This study has been presented in AUA21 (Virtual meeting, September 10-13, 2021, ref: 501059A).*

Study title	*P-score in pre-operative biopsies accurately predicts P-score in final pathology at radical prostatectomy in patients with localized prostate cancer (proposed title can be changed during publication)*
Author	Pontus Röbeck
Journal & Publication year	*Prostate, 2023*
DOI	doi: 10.1002/pros.24523.
AIM	*To assess the performance of the product by evaluating the concordance of the P-score in paired core needle biopsy (CNB) and radical prostatectomy (RP) specimens as well as in index vs. concomitant non-index tumor foci from the same RP. The rationale behind it was that intra-tumoral heterogeneity causes relatively low reliability of CNBs in determining the histological and genetic characteristics of a tumor.*
Cohort	*100 patients with localized prostate cancer, all patients were diagnosed by CNB and underwent RP between 2015 and 2018, Swedish Cohort.*
Method	*Gene expression was assessed with the Prostatype® RT-qPCR kit and the P-score was calculated. Patients were categorized into three P-score risk groups according to previously defined cut-offs.*
Results	*The CNB-based P-score groups were in substantial agreement with the RP-based P-score groups (weighted kappa score: 0.76 [95% CI 0.60-0.92]; Spearman's Rank Correlation coefficient r=0.83 [95% CI 0.74-0.89]; p<0.0001). Similarly, the P-score groups based on paired index tumor and concomitant non-index tumor foci (n=64) were also in substantial agreement (weighted kappa score: 0.74 [95% CI 0.57-0.91]; r=0.83 [95% CI 0.73-0.89], p<0.0001). Thus, the study provides evidence that the P-score obtained from readily available CNBs is representative of a patient's prostate tumor, and therefore constitutes a valuable treatment decision support tool. These findings underline the clinical value of the P-score which can add meaningful prognostic information and guide management of untreated PCa patients. This study has been presented in AUA22 (New Orleans, US, May 13-16, 2022, Ref: 501063A) and NUF22 (Helsinki, Finland, June 8-11, 2022, Ref: 501065A).*

External scientific validity findings for genomic biomarker applied in

(a) Prostatype® RT-qPCR kit

Evidence was found externally to support the prognostic value of the three genes used in our products in pre-clinical stage as well as in the last 5 years. The three genes' roles in cancer prognosis are described in detail below:

IGFBP3/IGFBP-3 (Insulin Like growth factor binding protein 3):

Pre-clinical stage (2005-2016)

The insulin-like growth factor binding protein IGFBP3's role as anti-tumor [4,5] or pro-tumor [6–8] in PCa both have supporting evidence. For example, Meht et al. reported that IGFBP3 knock-out (KO) mice failed to undergo apoptosis after castration. 80 weeks after crossing with a c-Myc-driven model, IGFBP3 KO mice tended to exhibit larger prostate tumors than the control mice and 55% of which were observed to have lung metastasis, indicating a metastasis-suppressing function of IGFBP3[5]. Meanwhile, in Severi et al.'s study, higher levels of IGFBP3 protein were associated with a higher level of PCa risk (p=0.008) in 605 PCa cases[7]. The association between IGFBP3 and PCa risk was explored and discussed in a meta-analysis including about 7,500 PCa cases. In that study, an inconsistent association between IGFBP3 and PCa risk among different **studies** was reported, which the authors concluded can be partly explained by study design: the magnitude of associations was smaller in prospective versus retrospective studies, and for IGFBP-3 the inverse association with PCa risk was seen in retrospective but not prospective studies[9].

Updated information in recent 5 years (2018-2023)

Cai *et al*. has reviewed comprehensively the anti-tumor and anti-metastatic role of IGFBP3 in multiple cancer including PCa. By extracting and analysing RNA-seq data from the TCGA database, they illustrated that higher expression of IGFBP3 correlated with better survival outcome in PCa which goes along with our findings. In addition to the previously known anti-tumor mechanism of IGFBP3 such as suppressing tumor-induced NF-kB activity, this review focused on a novel signaling pathway of IGFBP3 by activating caspase-8-induced apoptosis [10].

On the contrary, in another study conducted by Arai's group, they applied information extracted from the TCGA database and found high expression of IGFBP3 correlated with poor prognosis in PCa patients[11]. The distinct characteristics of the two cohorts from Arai and Cai's studies and different analysis methods might account for the opposite prognostic value of IGFBP3. Of note, IGFBP3 is cleaved, regulated, and degraded by different specific proteases in different tissue types. This may explain why IGFBP3 inhibits tumor proliferation in some cancer types like PCa and breast cancer while promoting tumor proliferation in other cancer types such as kidney or Glioma [10]. In castration-resistant prostate cancer (CRPC) patients particularly, Hensley *et al*., found that IGFBP3 correlated with tumor recurrence and poor patient survival, indicating a potential therapeutic value of IGFBP3 [12]. Aligned with Hensley et al.'s finding, Chen's group also

showed that IGFBP3 may promote the growth and migration of CRPC cells, and this was probably due to the deregulation of lysyl oxidase (LOX) [13].

Overall, the significant role of IGFBP3 in PCa has made it a popular biomarker to be used in assessing treatment or supplement effects, among which, different medicines such as bicalutamide, goserelin, leuprolide acetate, metformin hydrochloride, or supplement like pomegranate-extract were studied [14–16].

F3/tissue factor/TF:

Pre-clinical stage (2005-2016)

Tissue factor (TF/F3) has been implicated in promoting angiogenesis, tumor progression, and metastasis in a variety of tumors including PCa [17–21]. Its expression levels were also found to correlate significantly with Gleason score, stage of the disease in prostate specimens [18]. Strijbos *et al*. ran an investigation on serum markers in 162 CRPC patients treated with docetaxel-based therapy. They found circulating endothelial cells, circulating tumor cells, and TF levels alone and combined can predict early on overall survival (OS) [22]. Heinrich *et al*. conducted a study by using TF deficient prostate adenocarcinoma cells administered rat and found they exhibited significantly less metastasis and smaller tumor than the control rat. Reduction of tissue factor expression as they concluded may serve as immunotherapeutic strategy for preventing, and possibly, curing PCa [23].

Updated information in recent 5 years (2018-2023)

The widely recognized strong correlation between cancer and TF/F3 mediated pro-coagulation and thrombosis has been well reviewed, making TF a prognostic biomarker for the survival of cancer patients [24–27]. High levels of TF expression in tumors has shown to be associated with poor prognosis in a variety of cancers [25]. TF is also an FDA-approved drug target and the first TF-targeted antibody conjugate drug named Tisotumab Vedotin (HuMax®-TF-ADC) was approved by FDA in October 2021 [28]. This approval is based on the results of the single-arm, multicenter pivotal phase II innovaTV 204 study (NCT03438396), The results showed significant efficacy of Tivdak as a monotherapy, providing clinically meaningful and durable objective remissions: an overall remission rate (ORR) of 24%, median duration of remission (DOR) of 8.3 months, and a manageable safety profile [29]. Although the study was done on cervical cancer, it opens the possibility of targeting TF in other solid tumor types which express TF including PCa. Indeed, several clinical studies[30–32] are ongoing to investigate the safety and efficacy of continued treatment with HuMax®-TF-ADC in patients including PCa patients.

VGLL3 (Vestigial-like family member 3):

Updated information in recent 5 years (2018-2023)

Studies around role of VGLL3 in cancer is limited before 2016, only one study mentioned VGLL3 ‘s tumor suppressing role in ovarian cancer [33]. More scientific evidence emerged recently to support VGLL3 as a pro-tumor gene, which is in line with our scientific finding.

In Yamaguchi's review, it was summarized that the elevated expression of VGLL1-3 was observed in various types of tumors, and VGLL1-3 showed to possess tumorigenic functions. VGLL3 was involved in the activation of the MAPK, JAK-STAT, and WNT pathways together with enhanced immune infiltrate in cancer [34]. The role of VGLL3 in immune response was supported by Ren *et al*., who revealed VGLL3 as part of the immune-related risk score signature and was related to PCa patient's prognosis[35].

In addition to the pathways described above, VGLL3 was widely recognized to trigger several pathways by binding to its co-factor TEA domain-containing transcription factors (TEAD). Hori *et al*. found that VGLL3 promoted cell proliferation by activating the Hippo pathway [36], and they further discovered that the VGLL3-TEAD4 complex promotes EMT- like cell motility and therefore accelerates cancer progression [37].

Furthermore, VGLL3 was demonstrated by Zhang *et al*., to be a novel prognostic biomarker for gastric cancer, and VGLL3 was very likely involved in the Hippo pathway. The high expression VGLL3 group had a worse prognosis [38]. The same group later showed that VGLL3 was an independent unfavorable novel prognostic biomarker in stomach cancer patients. VGLL3 expression levels were significantly associated with poor differentiation, advanced Tumor-stage, and TNM stage [39].

External scientific validity for clinical parameters applied in P-score

P-score includes clinical parameters (PSA, cT-stage, & Gleason score), which are known standard diagnostic tools for PCa. In addition to their diagnostic roles, the combination usage of PSA, cT-stage, & Gleason score could serve as a risk stratification (prognosis) tool[40–42] and have well incorporated into clinical practice according to the clinical guidance worldwide, although variations exist in combination methods among different guidelines [43–45]. Notably, Gleason score, which later has evolved into new WHO/ISUP grading ranging from 1 to 5 [46,47] has shown great prognostic value [48–50]. We have investigated P-score before and after WHO/ISUP revision, P-score exhibits similar PCa specific prediction power (Saemundsson et al., manuscript in submission).

Analytical performance of the

(a) Prostatype® RT-qPCR kit

(b) P-score web service (PWS)

Analytical performance assesses the following items, depending on what is applicable for the product.

- Analytical sensitivity
- Analytical specificity
- Trueness (bias)
- Precision (repeatability and reproducibility)
- Accuracy
- Robustness
- Limits of detection and quantitation
- Measurement range
- Linearity
- Determination of appropriate criteria for specimen collection and handling, and control of known relevant endogenous and exogenous interference, cross-reactions.

(a) Analytical performance of Prostatype® RT-qPCR® kit

Performance characteristics	**Set up**	**Conclusion**
Analytical sensitivity	Samples with target (F3, IGFPB3, VGLL3) to GAPDH reference gene concentrations of 1 :50, 1 :100 and 1 :200 will be prepared for at least 6 PCRs. GAPDH target as background will be adjusted to a concentration of a typical biological sample (e.g., 4*104 copies/ µl which refers approximately to Cp 24). Respective target concentration of F3, VGLL3 and IGFBP3 then will be 800 (1 :50), 400 (1 :100), and 200 (1 :200) copies/µl.	**All dCp values were smaller than 7.7 within the acceptance criteria.** For each of the four assays, all 6 PCR wells showed an amplification curve (100% detection), and the median dCp values for the three target assays were 5.4 (F3), 6.3 (IGFBP3), and 6.8 (VGLL3). Even for the sample containing a 1:200 proportion between GAPDH and target assays all 6 x 4 amplification curves were present confirming a relative sensitivity for the three target assays below 1%. The limit of blank corresponds to the maximal observable Cp value of 38 for those genes. For GAPDH 16 of the 21 PCR wells showed an amplification curve with Cp values between 35.2 and 38.0. The limit of blank for GAPDH was estimated by the 95% quantile and equals Cp 35.2
Specificity Absence of interferences (accuracy)	The following potential interfering substances (concentration) were carefully selected and examined: Hematoxylin/Eosin (HE, 0.01% v/v), Xylol/Xylene (0.01% v/v), DNase I (0.06 U/µL), Human genomic DNA (20 ng/µL), Guanidinium Chloride (5 mM), Ethanol (1% v/v) and EDTA (1mM). Each sample will be measured parallel without any interfering substance (reference) and with added	None of the interfering substances does interfere with the Prostatype RT-qPCR kit at the concentration tested. Median Cp(substance) – median Cp(reference) for each target gene (IGFBP3, F3, VGLL3) and the reference gene GAPDH fell into the accepted criteria (≤1.0 and ≥-0.1)

	interfering substances mentioned above on the same plate. Aliquots from three RNA pools unspiked and spiked with potentially interfering substances were processed in two PCR runs.	
Precision (repeatability and reproducibility)	Reproducibility was evaluated by repeated testing of aliquots from eight RNA pools extracted from formalin fixed paraffin-embedded (FFPE) human prostate core needle biopsy containing cancer cells. Testing was performed by four operators at three different testing laboratories across Europe (Wild lab, Germany; Epigenomics, Germany; and Prostatype Genomics, Sweden).	The observed SD and CV are very low and within acceptance criteria, i.e., SD < 1.5 Cp, CV < 5%. The SD were ranged from 0.15-0.61, and the CV was ranged from 0.0063-0.02060.
	Compare the variations of gene expression levels between Unilabs Skövde and Prostatype Genomics using the same block from the same patient. A total of 34 Patient samples were analyzed, among which, (1) three patient samples were analyzed in two separate tubes at each site in order to confirm that RNA extraction yields reproducible results (intra-lab /intra-operator	For RNA extraction reproducibility (intra-lab correlation) assessment, results showed variations of ΔCp values < 1.5 for all genes and samples which is within the acceptance criteria. For inter-lab correlation assessment, results showed the average variation of ΔCp values was 0.98 ± 0.47 for IGFBP3, 0.92 ± 0.50 for F3 and 1.55 ± 1.43 for VGLL3. **In summary, ΔCp(IGFBP3) and ΔCp(F3) values were <1.5 and for the majority of samples and <2 for all samples.**

	correlation). (2) The remaining 31 patient samples were used to evaluate correlation between different sites and operators (inter-lab/inter-operator correlation). 8 sections derived from the same block and the same patient were divided into 2 batches of 4 sections in a randomized manner, tested by either Unilabs or PG. Each testing site prepared 12 independent PCR plates involving three different Prostatype® RT-qPCR Kit lots.	
Precision and reproducibility (Stability of the kit)	Long term stability assessment (shelf life) for Prostatype RT-qPCT Kit within a time frame up to 23 months. Five different manufacturing lots were tested. Three of the lots were followed up to 12 months, one was followed up to 23 months and one was followed up to 25 months.	All three genes were in specification in all five lots with only one exception. VGLL3 from one lot is out of specification after 12 months. The positive and negative controls were stable after 23 months. The master mix was stable after 23 months.
	Accelerated stability test (in-use simulation for PCR efficiency assessment). PCR was conducted at three time points. Within each run, negative and positive controls of the kit exposed to the in-use simulation.	**The calculated PCR efficiencies were all within the acceptance criteria.**
Precision and reproducibility	The expression levels differences of genes (GAPDH, IGFBP3, F3,	From the time when embedding fresh cells in FFPE blocks, mRNA degradation does occur to a different extent in the four genes.

(Stability of the RNA in stored specimen)	VGLL3) from the same FFPE blocks containing PC3 cells and MDA-MB-231 cells, were measured using the Prostatype RT-qPCR Kit within a time frame up to 3 months.	However, from the time when the cells are embedded in FFPE blocks, only minor degradation occurs. The STDEV of ΔCp for all three genes <1, within the deviation range of the Prostatype RT-qPCR Kit. **Thus, gene expression levels measured by the Prostatype RT-qPCR Kit are considerably stable and the time factor only has a minor influence on RNA quality in FFPE blocks.**
Robustness	PCR efficiency was determined by the primers and probes concentration optimization	**It was shown that the Prostatype RT-qPCR Kit was really robust toward changes of oligonucleotide concentrations.** Within the tested range no statistically significant performance influence could be observed.
	Ten simulated failure modes were investigated which have been reported by operators or are based on experience from internal use, such as wrong PCR analysis setting, pipetting error, or wrong storage of components.	**None of the tested failure modes turned out to be undetected, but critical (Type II).** The failure modes were either detected by the Prostatype RT-qPCR Control or did not critically impact the assay result.
Limit of Quantitation/Limit of Detection/Limit of Blank	Seven different target concentrations will be tested. All four genes will be mixed equimolar to determine the analytical performance of the multiplexed RT-qPCR. Highest target concentration will be $2.5*10^6$ copies (1:1) RNA-construct of each gene per microliter sample. Lower concentrated samples are tested containing the targets in following ratios: 1:4, 1:16, 1:64, 1:256, 1:1024, 1:4096; Three	**The observed SD and CV are low and within acceptance criteria, i.e., SD < 1.5 Cp, CV < 5%.** The largest SD of 0.44 was detected for VGLL3 at a dilution of 1:4096, and the largest CV was 1.5% for VGLL3 at a dilution of 1:64. The 1:4096 dilution refers to a target concentration of 6.1*102 copies. The data showed the four assays (genes) are quantitative at least down to this concentration. Hence, The Limit of Quantitation of the Prostatype® RT-qPCR refers to a target concentration of 6.1*102 copies. The lowest concentrations (Limit of Detection) showing amplification curves for all replicates (100% detection) were as follows: GAPDH: 38 copies/µL, F3: 38 copies/

	different verifications lots were proceeded. Limit of Quantitation is defined as the minimal concentration where the standard deviation of repeated measurement is smaller than 1.5 Cp and the CV is smaller than 5 %. Limit of Detection is defined as a concentration which represents a Cp value of 38 (maximal number of PCR runs). The Limit of Detection was determined by extrapolating the regression model to a Cp value of 38 and calculation of corresponding concentration. Apart from the initial seven target concentrations that was tested, another four lower concentrations were tested in 24 replicates (VER_0112R_rev1_AttA_Precision (table 9 and figure 6).	µL, IGFBP3: 76 copies/µL, VGLL3: 153 copies/µL. The limit of blank corresponds to the maximal observable Cp value of 38 for those genes. For GAPDH 16 of the 21 PCR wells showed an amplification curve with Cp values between 35.2 and 38.0. The limit of blank for GAPDH was estimated by the 95% quantile and equals Cp 35.2.
linearity	To establish the range of linearity for each of the four genes 7 different target concentrations will be tested. Target of all four genes will be mixed equimolar to determine the analytical performance of the multiplexed RT-qPCR. Highest target concentration will be 2.5*106 copies (1 :1) RNA-construct of each gene per	For the linear model the observed coefficient of determination R^2 was 0.997 for GAPDH, IGFPB3, F3 and 0.995 for VGLL3. These results indicate that over the range of tested concentrations the relationship between assay results (CP value) and RNA input (log transformed) is well describes by a linear regression model for each of the four genes.

	microliter sample. Lower concentrated samples are tested containing the targets in following ratios: 1 :4, 1: 16, 1 :64, 1 :256, 1 :1024, 1 :4096; resulting in a lowest target concentration of 610 target copies per micro liter sample For each concentration 15 PCRs will be performed in 3 runs, i.e in total 105 PCR wells will be measured. Regression models with the Cp values as response and concentration as predictors will be compared to assess the range of linearity. The range of linearity was defined by the largest subset of subsequent diluted samples where no nonlinear (e.g. quadratic) effect exists. Linearity was evaluated by testing a dilution series of synthetic target RNA covering a range from $2.5x10^6$ to 38 molecules/ µL. Testing comprised three different Prostatype® RT-qPCR Kit Lots. For GAPDH and F3 assay the linear range covers the concentrations from $2.5x10^6$ to 38 copies/ µL. For IGFBP3 linearity is given down to 153 copies/ µL, and for VGLL3 down to 305 copies/µL.	

RNA input and tissue area	All patients samples were from patients over 18 years old with localized prostate cancer, hence from the intended patient population. 1) There were two rounds of experiments with comparisons of regular tissue amounts (10-20 mm2) and big chunks (30-60 mm2). The first round included 18 samples from 9 patients, where each patient had one sample with regular tissue amount and one big chunk sample. The big chunk samples contained approximately 50%-67% benign tissue. After analysis of the first round, a second round of experiments was conducted where a higher proportion of tumor tissue (approximately 2/3 tumor 1/3 benign) was included in the big chunk sample. The second round included 13 individuals, 9 of which one big chunk sample was compared to a single regular tissue sample from the same patient, and 4 individuals where one big chunk sample was compared to two regular tissue samples from the same patient. Expression levels of IGFBP3, F3, VGLL3 and GAPDH were measured using a pre-production	1) In round 1, the GAPDH Ct value was similar to or higher in the big chunk compared to the regular tissue. This indicates that a too large amount of benign tissue was included for Prostatype analysis. In the second round of experiments, the average improvement (reduction) Ct values is 1.8 Ct units. This means that the big chunk samples should contain 2^1.8 = 3.48 times more material. Increasing the input from 10-20 mm2 to 30-60 mm2 increased the success rate from 76% to 100%. 2) For samples with a tumor area between 15mm^2 and 30mm^2 and 50% cancer cells, there was >90% success rate to obtain Cp(GAPDH) ≤28 from these specimens. For specimens with a tumor area 5-15mm^2 and 50% cancer cells, there was around 80% success rate to generate Cp(GAPDH) ≤28. Samples with <5mm^2 tumor area are not suitable for testing with Prostatype® RT-qPCR Kit. For newly diagnosed patients' FFPE core needle biopsy samples collected in the whole Sweden (31 patients, <18months prepared), Maxwell Promega kit has performed 100% successful rate so far. The study showed that in practical clinical settings, Prostatype test performs much better for newly prepared FFPE core needle biopsy samples collected from newly diagnosed patients.

	version of the commercial Prostatype RT-qPCR kit. 2) The amount of tissue in mm^2 (<$5mm^2$ tumor area - was also investigated in another report. We performed technical data analysis for these samples such as drop-out rate analysis in relation to tissue input.	

(b) Analytical performance of P-Score web service (PWS)

The expression levels of IGFBP3, F3, and VGLL3 genes in tumour biopsies and PSA, Gleason score (cGleasonscore stands for clinical Gleason score), and tumour stage (cTstage stands for clinical tumour stage) as clinical parameters.

Data: merged data from Stockholm-Uppsala (Sweden), Malmö (Sweden) and Taiwan datasets. After cleaning in total, it is 1004 patients' data.

	IGFBP3	F3	VGLL3	PSA	cGleasonscore	cTstage
count	1004.000000	1004.000000	1004.000000	1004.000000	1004.000000	1004.000000
mean	2.656265	2.134855	6.036773	34.664771	7.185259	1.910359
std	1.295983	1.840636	3.349588	125.969197	1.028770	0.849569
min	-1.390000	-2.110000	-1.190000	0.500000	6.000000	0.000000
25%	1.860000	0.947500	3.340000	6.200000	6.000000	1.000000
50%	2.640000	2.025000	6.430000	9.900000	7.000000	2.000000
75%	3.355000	3.080000	8.550000	23.000000	8.000000	3.000000
max	10.060000	10.630000	14.040000	2749.980000	10.000000	4.000000

Simulated data: Some methods required simulation data and the ways data was simulated is described below.

Analytical Sensitivity

Objective: To determine if the software can detect and accurately reflect small but significant changes in gene expression levels and clinical parameters.

For each pair of consecutive data points, the difference in the P-score (delta_pscore) is calculated and the difference in the parameter value (delta_param). Then, delta_pscore is divided by delta_param for each pair and an average is calculated across all pairs, leading to an average rate of change. Average Sensitivity Index is similar to the Average Rate of Change, but we take the absolute value of the rate of change before averaging.

Interpretation: A high Average Rate of Change indicates that the output is very responsive or sensitive to changes in the input parameter. A positive value suggests a direct relationship (P-score increases as the parameter increases), while a negative value indicates an inverse relationship. Average rate of change is a measure that quantifies how much the output (P-score) changes on average for a unit change in an input parameter. The large rate of change and high sensitivity index for parameters like the Gleason score, cTstage and PSA are aligned well with

clinical reality, where crossing certain thresholds significantly alters the patient's prognosis or treatment strategy.

Parameters	Average rate of change	Average sensitivity index
IGFBP3	-49851513.46	229315261.89
F3	0.47	6.99
VGLL3	0.44	5.89
PSA	39880359.18	179461615.97
cGleasonscore*	538384847.29	8175473581.34
cTstage*	378863411.40	9790628117.86

*clinical Gleasonscore (cGleasonscore), clinical Tstage (cTstage)

IGFBP3 – has large negative average rate of change, indicating that the P-score decreases significantly as IGFBP3 increases. The magnitude of change suggests that IGFBP3 has a substantial inverse relationship with the P-score. A small increase in IGFBP3 expression leads to a large decrease in the P-score. The sensitivity index's high value reinforces the notion that IGFBP3 is a highly influential parameter in determining the P-score. It’s very sensitive to changes in IGFBP3 and even small changes in IGFBP3 result in large changes in the P-score.

Both F3 and VGLL3 show moderate average rate of change. Much smaller, positive values indicate a direct relationship with the P-score, but the impact of changes in these genes’ expressions are much more moderate compared to IGFPB3. This suggests that while F3 and VGLL3 influence the P-score their effect are not as pronounced. Both F3 and VGLL3 have also reasonable average sensitivity index, implying a noticeable but not overwhelming sensitivity of the P-score to changes in these genes’ expressions.

PSA in contrast to IGFBP3 has large positive average rate of change, indicating that the P-score increases significantly as PSA increases. The magnitude of change suggests that PSA has a substantial direct relationship with P-score. A small increase in PSA leads to large increase in P-score. The sensitivity index confirms that PSA is highly influential parameter in determining P-score.

Both cGleasonscore and cTstage have largest positive average rates of change, indicating that the P-score increases significantly as these parameters increase. The highest magnitude of change suggest that cGleasonscore and cTstage have strongest direct relationship with P-score. A small increase in any of these parameters lead to large increase in P-score. The highest sensitivity index confirms that cGleasonscore and cTstage have the strongest influence on P-score.

Statistical tests around thresholds

Testing whether P-score significantly changes when a parameter crossing a threshold.

Determination whether parametrical or non-parametrical statistics shall be applied:

None of the used parameters to calculate P-score follow the normal distribution (neither P-score for that matter). Therefore non-parametrical statistic shall be applied.

When a parameter has only one threshold, Mann-Whitney U test test can be applied:

IGFBP3: Mann-Whitney U test for parameter at threshold: U-Statistic=209987.0, P-value=2.16e-77

F3: Mann-Whitney U test for parameter at threshold: U-Statistic=58973.0, P-value=7.66e-23

VGLL3: Mann-Whitney U test for parameter at threshold: U-Statistic=87917.0, P-value=7.05e-16

When a parameter has multiple thresholds, Kruskal-Wallis test is applied at first and then post-hoc Dunn test with Bonferroni adjustment:

PSA: 4 thresholds

Kruskal-Wallis H test: H-Statistic=666.56, p-value=6.05e-143

It means that there is a statistically significant differences in the median P-score among data groups at the different sides of thresholds.

	1	2	3	4	5
1	1.000000e+00	4.105348e-18	7.102783e-45	1.727939e-46	3.629108e-126
2	4.105348e-18	1.000000e+00	1.420628e-08	3.422054e-17	1.145871e-58
3	7.102783e-45	1.420628e-08	1.000000e+00	1.389656e-04	1.798061e-21
4	1.727939e-46	3.422054e-17	1.389656e-04	1.000000e+00	5.866222e-02
5	3.629108e-126	1.145871e-58	1.798061e-21	5.866222e-02	1.000000e+00

The transitions between thresholds are all statistically significant ($p < 0.05$) except the boundary between group 4 and 5 (last threshold) which has $p = 0.057$.

cTstage: 2 thresholds

Kruskal-Wallis H test: H-Statistic=345.93, p-value=7.61e-76

Dunn test with Bonferroni adjustment:

	1	2	3
1	1.000000e+00	4.365768e-61	2.883728e-23
2	4.365768e-61	1.000000e+00	1.039459e-02
3	2.883728e-23	1.039459e-02	1.000000e+00

The transitions between all the thresholds are statistically significant.

cGleasonscore: 3 thresholds Kruskal-Wallis H test: H-Statistic=286.42, p-value=6.37e-63

Dunn test with Bonferroni adjustment:

	1	2	3	4
1	1.000000e+00	3.800708e-24	1.488447e-46	1.074407e-106
2	3.800708e-24	1.000000e+00	3.706780e-13	3.150673e-52
3	1.488447e-46	3.706780e-13	1.000000e+00	4.790517e-08
4	1.074407e-106	3.150673e-52	4.790517e-08	1.000000e+00

The transitions between all the thresholds are statistically significant.

Statistical Significance: The significant p-values indicate that there are statistically significant differences in the median P-scores between tested groups of data. This means that the changes

in the parameter that define these group boundaries are associated with substantial changes in the P-score.

Clinical Interpretation: Since these groups of data are determined based on clinically relevant thresholds, this finding suggests that crossing practically all these thresholds has a meaningful impact on the P-score, which could translate to significant clinical implications.

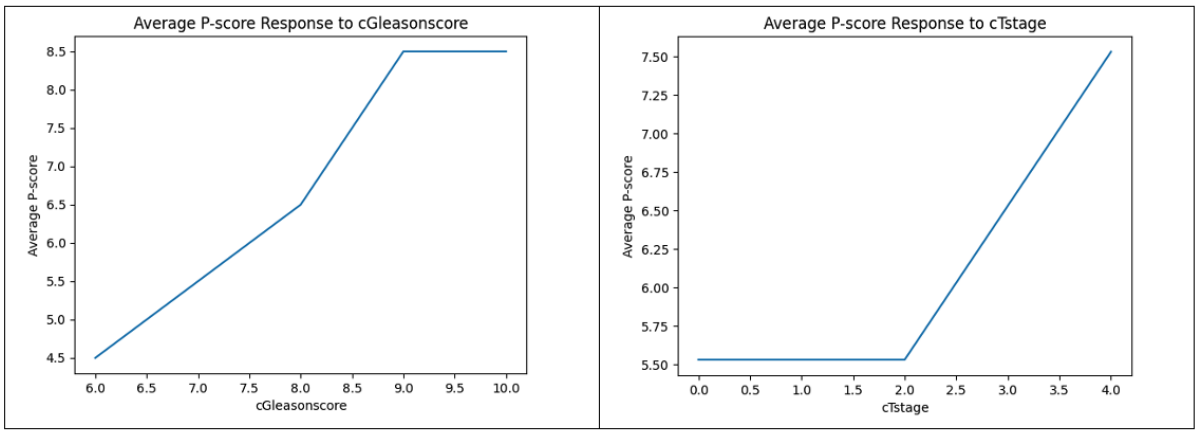
The only one tested threshold (the 4th of PSA) was failed short of reaching statistical significance. The non-significance of a threshold in the analysis could indicate that, within the range of tested data and the specific combinations of parameter values present, crossing this threshold does not consistently lead to significant changes in the P-score. It suggests that the model's sensitivity to changes in this particular parameter, at least at this specific threshold, might be lower than expected. This could be due to the interplay of other parameters in the model.

Analytical Specificity: Testing for influence of unrelated data

Two datasets were used; the original dataset and the other one where original data were contaminated with 5 variables representing random noise. Both datasets were fed into the algorithm and P-scores generated by both datasets were compared. The P-scores were compared and were found identical. This result indicates that the added noise did not alter the software's output, suggesting a high degree of analytical specificity.

Trueness (Bias)

The Internal Consistency Check was used to evaluate trueness/bias. This technique involves analysing how the P-score responds to systematic variations in key input parameters, specifically well-established and clinically significant parameters like the Gleason score and T-stage. The average P-score is plotted against a range of values for the clinical Gleasonscore (cGleasonscore) and clinical Tstage (cTstage). Typically, higher Gleason scores and Tstage are associated with more aggressive disease and worse prognosis. Each point on these graphs represents the average P-score calculated for all instances in the dataset with a specific cGleasonscore or cTstage. The calculation is done by systematically altering the value of the cGleasonscore or cTstage across its entire range, one value at a time, while keeping other parameters constant. The Internal Consistency Check thus serves as a crucial part of the analytical performance evaluation for our P-score model. It provides a comprehensive assessment of how well the model captures and reflects the nuances of clinically important parameters, ensuring its reliability and appropriateness for clinical applications.



cGleasonscore

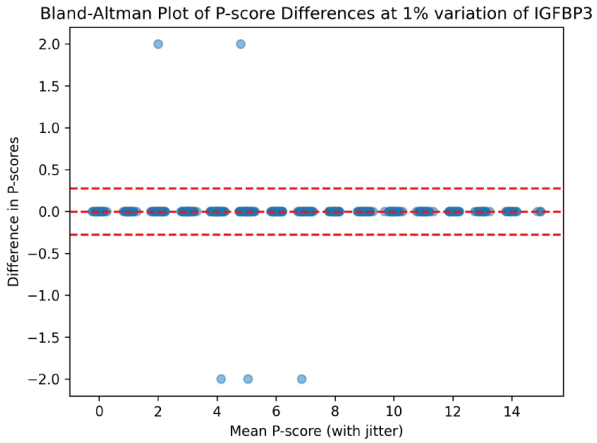
The break in the line at Gleason score 8 suggests a significant change in the model's risk assessment at this point. This is due to a threshold in the model where the Gleason score crosses a critical value, leading to a different weight being applied. The line levelling off from Gleason score 9 onwards indicates that increases in the Gleason score beyond this point do not substantially change the P-score. This could mean that the model considers Gleason scores above 9 to represent a similar level of risk, and it is a ceiling effect in how the model is constructed.

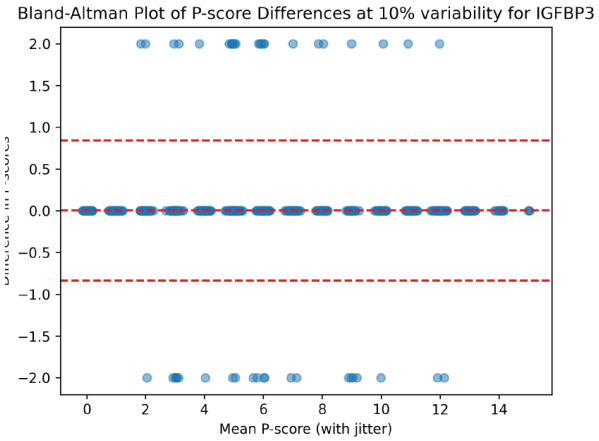
cTstage

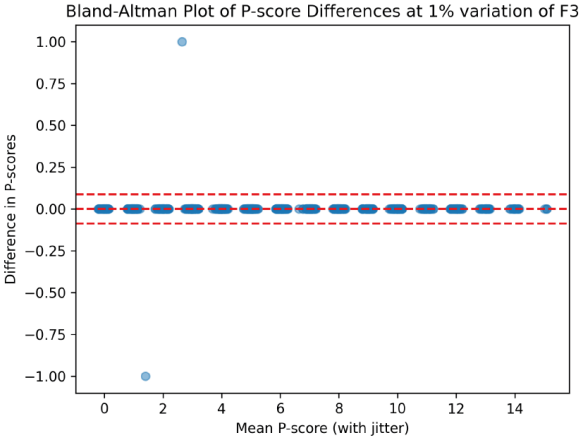
Break at cTstage = 2: The break in the line at cTstage suggests a significant change in the model's risk assessment at this point. This is due to a threshold in the model where the cTstage crosses a critical value, leading to a different weight being applied.

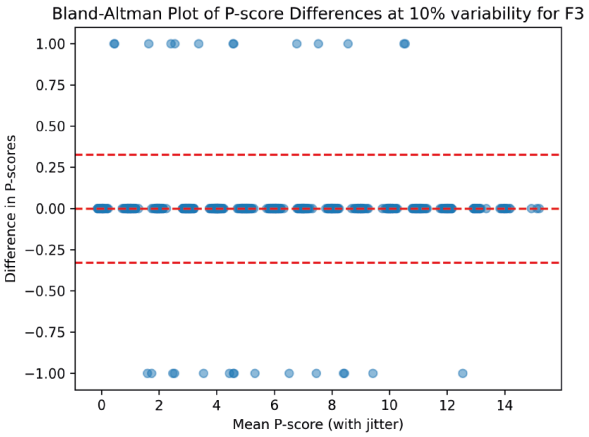
Reproducibility/ Model Robustness

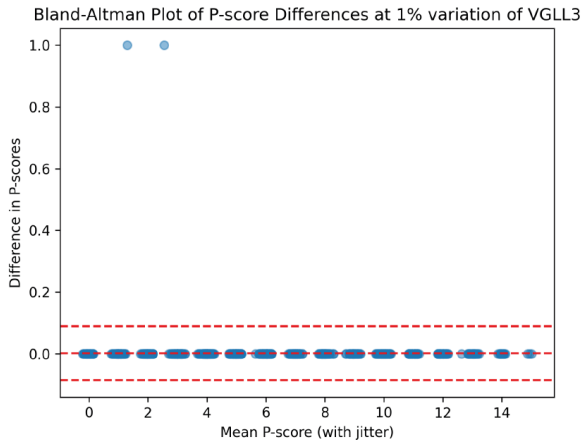
The robustness of the algorithm is assessed by introducing variation in the input parameters one by one and assessing the impact on P-score with a help of Bland-Altman plot. The variation for continuous variables IGFBP3, F3, VGLL3 and PSA was applied as random changes +/-1% and +/-10%. For the ordinal variables, cGleasonscore and cTstage, the random changes were applied as +/- 1 unit under condition that each of these variables will remain withing its range in the dataset. The range for cGleasonscore was 6-10 and for cTstage 0-4. Jitter and transparency were applied to datapoints in order to identify clustering and patterns. The limits of agreement, two dashed lines below and above complete agreement around the zero point are defined as mean difference +/- 1.96 x STD. Data clustering observed above and below the complete agreement line in our Bland-Altman plots is indicative of the passage of certain thresholds for each parameter, leading to increases or decreases in P-score. This threshold effect is particularly pronounced in our analysis of continuous input parameters like IGFBP3, F3, VGLL3, and PSA, where subtle variation (1%) results in only a marginal increase in data spread. However, with stronger variation (10%), the spread becomes substantially more pronounced, highlighting the model's increased sensitivity to larger changes in these parameters.

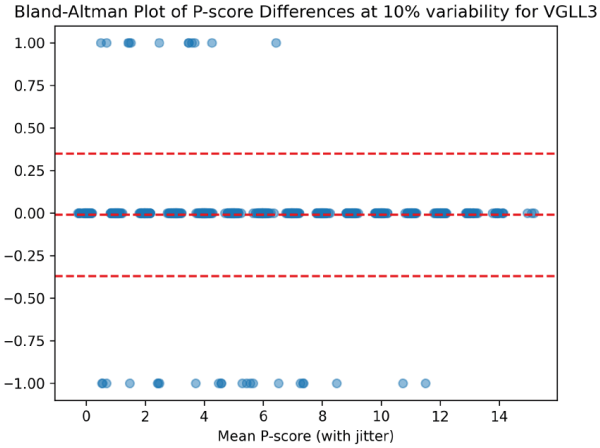

Bland-Altman Plot of P-score Differences at 1% variation of IGFBP3
Difference in P-scores
Mean P-score (with jitter)

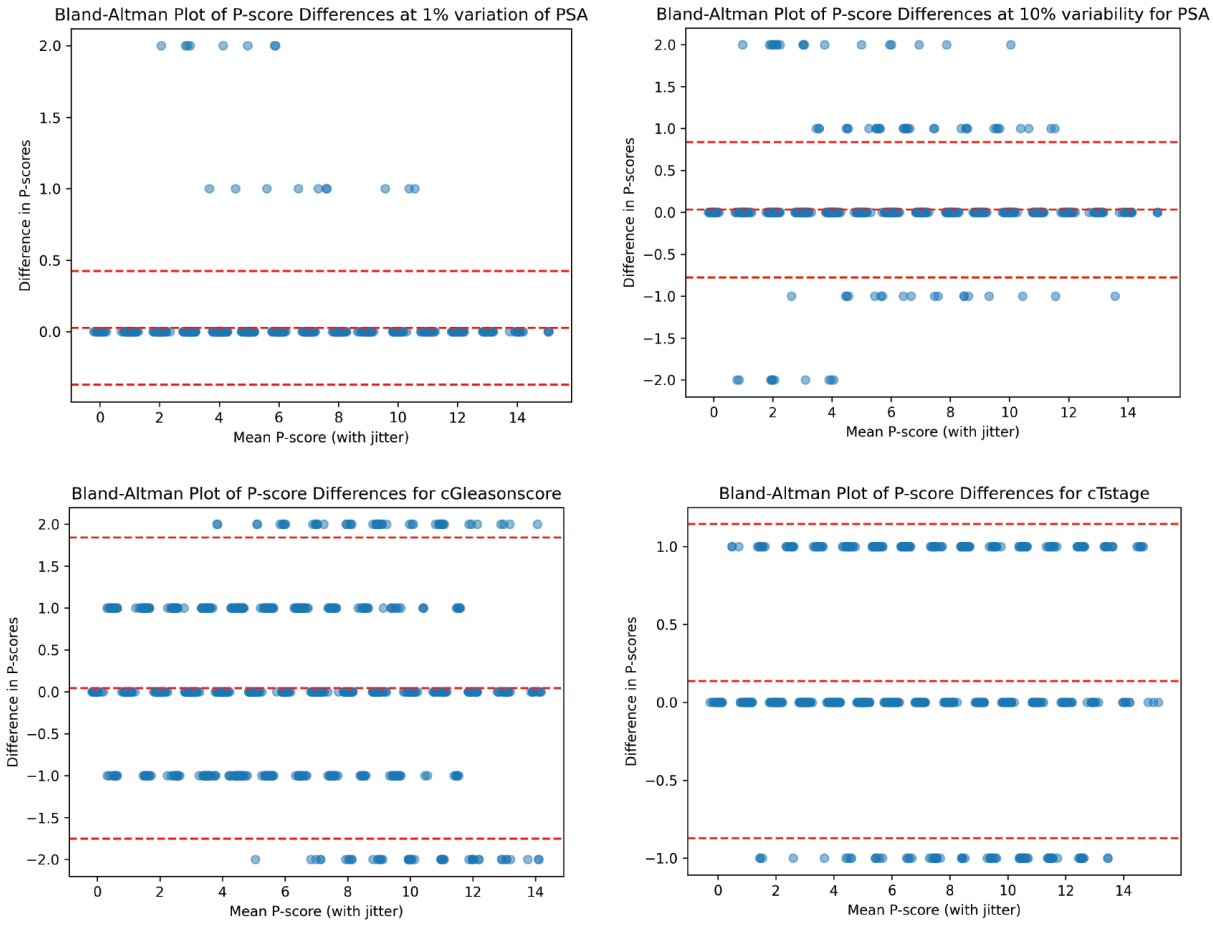

Bland-Altman Plot of P-score Differences at 10% variability for IGFBP3
Mean P-score (with jitter)


Bland-Altman Plot of P-score Differences at 1% variation of F3
Difference in P-scores
Mean P-score (with jitter)


Bland-Altman Plot of P-score Differences at 10% variability for F3
Difference in P-scores
Mean P-score (with jitter)


Bland-Altman Plot of P-score Differences at 1% variation of VGLL3
Difference in P-scores
Mean P-score (with jitter)


Bland-Altman Plot of P-score Differences at 10% variability for VGLL3
Mean P-score (with jitter)



Accuracy

We defined Low risk when cGleasonscore was 6 and less and cTstage was 2 and below; Medium risk defined when cGleasonscore was 7 or cTstage was 3, High risk defined when cGleasonscore was 8 and higher or cTstage was 4. These were our approximations of true outcomes, and we compared it with our categorized P-scores (P-score 0-2 was low risk group, 3-5 as medium risk group and above 5 as high-risk group). Mean square errors (MSE), measures the average squared difference between the estimated values (P-scores) and what is estimated as the true value; risk categories defined in our case by cGleasonscore and cTstage.

Mean square error (MSE) was 0.59

The value of 0.5946 indicates the average squared error in the P-score model's predictions relative to the clinically defined categories. In the context of MSE, a lower value is generally better, indicating smaller errors. MSE of 0.59 suggests that while the model's predictions are generally in line with the clinical categorization, there is still a notable average error in these predictions. To put it in perspective, consider the maximum possible squared error on our scale. For example, the worst-case error would be misclassifying a high-risk patient (true value = 3) as

low risk (predicted value = 1), resulting in a squared error of $(3 - 1)^2 = 4$. Compared to this, an average squared error of 0.59 is lower, but it still indicates room for improvement. MSE of 0.59 suggests that while the model's predictions are generally in line with the clinical categorization, there is still a notable average error in these predictions. It is important to assess where these errors are occurring and their potential impact on clinical decision-making, as well as to explore ways to further refine and improve the model.

Limits of Detection and Measurement Range

Limits of Detection: Establish the smallest change in input parameters that the software can detect reliably. This is done through a systematic sensitivity analysis to identify the minimal changes in inputs that the software can consistently detect, as indicated by the output score.

A realistic range of values for each parameter is defined based on our combined dataset that are clinically relevant; IGFBP3, F3, VGLL3, PSA, cGleasonscore and cTstage. Each parameter is adjusted iteratively across its range while other parameters are held constant. At each step, the output score is calculated using the software's algorithm to determine detection limits.

Measurement Range

The objective of the measurement range analysis is to establish the range of input values for which the software accurately calculates the risk score (P-score). This ensures that the software remains reliable and accurate across the spectrum of clinically relevant values.

For each parameter, iteratively set the parameter to its extreme, but possible values, while keeping other parameters at their typical levels. P-score range is established between 0 and 15. The extreme values for the input parameters did not result in P-score going beyond these limits. In the evaluation, it was observed that within the typical range for each parameter (e.g., IGFBP3, F3, VGLL3, PSA, cGleasonscore, cTstage), the software reliably produced P-scores that align with clinical expectations. Importantly, when testing the software with input values extending beyond these standard ranges, the P-scores continued to be consistent and within an interpretable range. This suggests that the software is robust even under atypical input scenarios.

P-score for each extreme scenario:

Input parameter	Extreme values	P-score
IGFPB3	-2	6
	15	4
F3	-2	5
	15	6
VGLL3	-2	5
	15	6
PSA	0	3
	5000	8
cGleasonscore	5	4
	11	8
cTstage	0	5
	5	7

VI.II Summary of performance data from the equivalent device, if applicable

Not an equivalent device.

VI.III Summary of performance data from conducted studies of the device prior to CE-marking

Specifically for companion diagnostics and not applicable for Prostatype RT-qPCR kit and PWS.

VI.IV Summary of performance data from other sources, if applicable

All the performance data are summarized under the 'Summary of scientific validity of the device' and 'Ongoing or planned post-market performance follow-up' sections.

VI.V An overall summary of the performance and safety

To conclude, Prostatype kit and PWS are safe to use for prognostic evaluation of prostate cancer. Clinical evidence showed good correlation between the genes and PCa prognosis as well as between the clinical parameters used and PCa. Data from the analytical tests and performance studies demonstrates that they are safe to use according to the intended purpose. The company will continue evaluating the performance of these devices and document them accordingly.

When it comes to clinical utility of the Prostatype® RT-qPCR kit and the PWS the balance between the benefit vs. risk needs to be considered. Prostatype test is not a diagnostic test but a prognostic test which could be used as a decision support tool for the urologist/ oncologist when they need a second opinion. The test result classifies tested patients in three different risk groups but does not give any advice or suggestion on which treatment to be chosen for the tested patient. It is always up to the treating doctor to choose a treatment path for a patient based on his clinical parameters, family history, age, and other risk factors. However, the clinical benefit of using the Prostatype® RT-qPCR kit and the PWS could be reassurance of their justification of addressing how bad the cancer is. Therefore, the risk of having serious consequences from using the product is justified very low compared to the benefit of using it.

VI.VI Ongoing or planned post-market performance follow-up

As a process of post-market surveillance, data from the tested clinical samples were analyzed to check the products performance. All the results undergo careful analysis to check if the data meets the acceptance and validity criteria of the kit according to the Instructions for use (IFU).

From first quarter of 2018 to 2022-11-10, patients were tested 37 times (total 52 patients since in some test more than one samples were tested) at the Prostatype Genomics clinical laboratory. In these 37 qPCR runs, Prostatype kits from four different batches were used. In all the tests the CT values of GAPDH in both the positive and negative control successfully met the acceptance and validity criteria (figure 1).

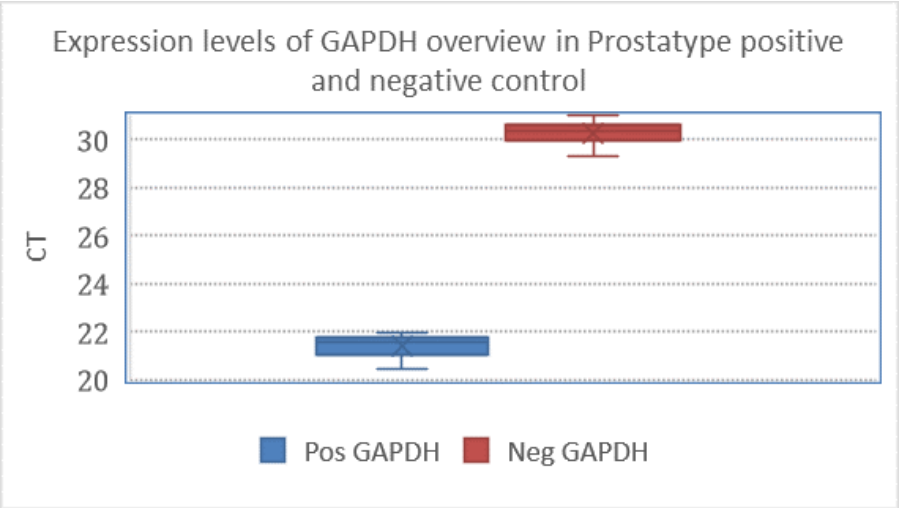


Figure 1 The plot shows distribution of GAPDH CT in the positive and negative controls of the kits used in the Prostatype Genomics clinical laboratoryfrom 2018- 2022-11-10. N=37 tests, and no. of patients tested were 52. Acceptance range for the Positive control GAPDH CT is 20.00 ≤ Median Cp ≤ 22.00; and negative control GAPDH CT is 29.00 ≤ Median Cp ≤ 31.00.

The average GAPDH CT value from these 37 positive control analyses was 21.40 with a standard deviation of 0.44 and for negative control analyses was 30.25 with a standard deviation of 0.48.

The ΔCT IGFBP3 (GAPDH-IGFBP3), ΔCT F3 (GAPDH-F3), were also met in all the tests. In only one of the cases of 37 the ΔCT VGLL3 (GAPDH-VGLL3) was deviated from acceptance criteria (Figure 2). Average ΔCT of IGFBP3 was 2.94 with a standard deviation of 0.40, of F3 was 2.26 with a standard deviation of 0.35, and of VGLL3 was 3.76 with a standard deviation of 0.49.

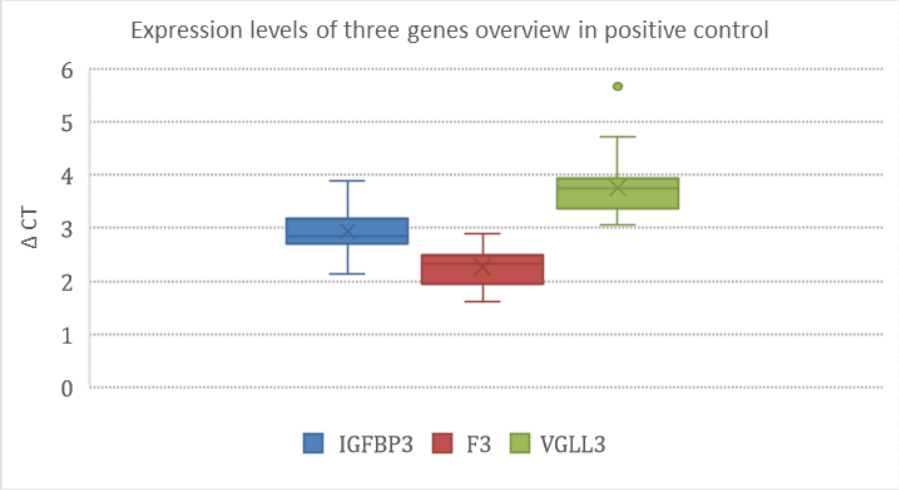


Figure 2 The plot shows distribution of median ΔCT for IGFBP3, F3, & VGLL3 of positive controls from the kits used in the Prostatype Genomics clinical laboratory from 2018- 2022-11-10. N=37 tests and no. of patients tested were 52. Acceptable median ΔCT ranges are: 1.90 ≤ IGFBP3 Median Delta Cp ≤ 3.90; 0.90 ≤ F3 Median Delta Cp ≤ 2.90; 2.80 ≤ VGLL3 Median Delta Cp ≤ 4.80.

Additionally, 58 Prostatype RT-qPCR kits (16x2 patient) were sold during 2019-2021 under the umbrella of a validation study that was performed externally. No negative feedback or complaints for the products were received.

VII. Metrological traceability of assigned values

No metrological assigned values are used in the products.

VII.I Explanation of the unit of measurement, if applicable

No metrological assigned values are used in the products.

VII.II Identification of applied reference materials and/or reference measurement procedures of higher order used by the manufacturer for the calibration of the device

No metrological assigned values are used in the products.

VII. Suggested profile and training for users

The Prostatype RT-qPCR kit is for use in a professional laboratory and only be used by professional laboratory personnel that has been trained and certified by Prostatype Genomics according to the corresponding training document.

The PWS is for use only by trained and certified laboratory staff or medically trained and certified healthcare professionals.

6. References

1. Peng, Z. *et al.* An expression signature at diagnosis to estimate prostate cancer patients' overall survival. *Prostate Cancer Prostatic Dis.* **17**, 81–90 (2014).

2. Peng, Z. *et al.* Improving the Prediction of Prostate Cancer Overall Survival by Supplementing Readily Available Clinical Data with Gene Expression Levels of IGFBP3 and F3 in Formalin-Fixed Paraffin Embedded Core Needle Biopsy Material. *PLOS ONE* **11**, e0145545 (2016).

3. Söderdahl, F., Xu, L.-D., Bring, J. & Häggman, M. A Novel Risk Score (P-score) Based on a Three-Gene Signature, for Estimating the Risk of Prostate Cancer-Specific Mortality. *Res.*

Rep. Urol. **Volume 14**, 203–217 (2022).

4. Chen, C. *et al.* Prostate carcinoma incidence in relation to prediagnostic circulating levels of insulin-like growth factor I, insulin-like growth factor binding protein 3, and insulin. *Cancer* **103**, 76–84 (2005).

5. Mehta, H. H. *et al.* IGFBP-3 is a metastasis suppression gene in prostate cancer. *Cancer Res.***71**, 5154–5163 (2011).

6. Platz, E. A. *et al.* Plasma insulin-like growth factor-1 and binding protein-3 and subsequent risk of prostate cancer in the PSA era. *Cancer Causes Control* **16**, 255–262 (2005).

7. Severi, G. *et al.* Circulating Insulin-Like Growth Factor-I and Binding Protein-3 and Risk of Prostate Cancer. *Cancer Epidemiol. Biomarkers Prev.* **15**, 1137–1141 (2006).

8. Rothermundt, C. *et al.* Metformin in chemotherapy-naive castration-resistant prostate cancer: a multicenter phase 2 trial (SAKK 08/09). *Eur. Urol.* **66**, 468–474 (2014).

9. Rowlands, M.-A. *et al.* Circulating insulin-like growth factor peptides and prostate cancer risk: A systematic review and meta-analysis. *Int. J. Cancer* **124**, 2416–2429 (2009).

10. Cai, Q., Dozmorov, M. & Oh, Y. IGFBP-3/IGFBP-3 Receptor System as an Anti-Tumor and Anti-Metastatic Signaling in Cancer. *Cells* **9**, E1261 (2020).

11. Arai, T. *et al.* Pirin: a potential novel therapeutic target for castration‐resistant prostate cancer regulated by *miR‐455‐5p*. *Mol. Oncol.* **13**, 322–337 (2019).

12. Hensley, P. J. *et al.* Predictive and targeting value of IGFBP-3 in therapeutically resistant prostate cancer. 16.

13. Chen, X. *et al.* Downregulation of LOX promotes castration-resistant prostate cancer progression via IGFBP3. *J. Cancer* **12**, 7349 (2021).

14. National Cancer Institute (NCI). *Phase II Study of Metformin in a Pre-prostatectomy Prostate Cancer Cohort*. https://clinicaltrials.gov/ct2/show/NCT01433913 (2017).

15. National Cancer Institute (NCI). *A Randomized Phase II Study of Androgen Deprivation Combined With IMC-A12 Versus Androgen Deprivation Alone for Patients With New Hormone Sensitive Metastatic Prostate Cancer*. https://clinicaltrials.gov/ct2/show/NCT01120236 (2018).

16. National Cancer Institute (NCI). *A Phase IIA Exploratory, Randomized, Placebo-Controlled Trial of Pomegranate Fruit Extract/Pomx™ in Subjects With Clinically Localized Prostate Cancer Undergoing Active Surveillance*. https://clinicaltrials.gov/ct2/show/NCT02095145 (2020).

17. Abdulkadir, S. A. *et al.* Tissue factor expression and angiogenesisin human prostate carcinoma. *Hum. Pathol.***31**, 443–447 (2000).

18. Kaushal, V. *et al.* Expression of Tissue Factor in Prostate Cancer Correlates With Malignant Phenotype. *Appl. Immunohistochem. Mol. Morphol.* **16**, 1–6 (2008).

19. Haubold, K. *et al.* Tissue factor procoagulant activity of plasma microparticles is increased in patients with early-stage prostate cancer. *Thromb. Haemost.* **101**, 1147–1155 (2009).

20. Lindholm, P. F. *et al.* Role of monocyte-lineage cells in prostate cancer cell invasion and tissue factor expression. *The Prostate* **70**, 1672–1682 (2010).

21. Hisada, Y., Geddings, J. E., Ay, C. & Mackman, N. Venous thrombosis and cancer: from mouse models to clinical trials. *J. Thromb. Haemost.* **13**, 1372–1382 (2015).

22. Strijbos, M. H. *et al.* Circulating endothelial cells, circulating tumour cells, tissue factor, endothelin-1 and overall survival in prostate cancer patients treated with docetaxel. *Eur. J. Cancer* **46**, 2027–2035 (2010).

23. Heinrich, J. E. *et al.* Vaccination against prostate cancer using a live tissue factor deficient cell line in Lobund–Wistar rats. *Cancer Immunol. Immunother.* **56**, 725–730 (2007).

24. Graf, C. & Ruf, W. Tissue factor as a mediator of coagulation and signaling in cancer and chronic inflammation. *Thromb. Res.* **164**, S143–S147 (2018).

25. Hisada, Y. & Mackman, N. Cancer cell-derived tissue factor positive extracellular vesicles: biomarkers of thrombosis and survival. *Curr. Opin. Hematol.* **26**, 349 (2019).

26. Li, H. *et al.* Tissue factor: a neglected role in cancer biology. *J. Thromb. Thrombolysis* **54**, 97–108 (2022).

27. Van Dreden, P., Elalamy, I. & Gerotziafas, G. T. The role of tissue factor in cancer-related hypercoagulability, tumor growth, angiogenesis and metastasis and future therapeutic strategies. *Crit. Rev. Oncog.* **22**, (2017).

28. FDA Approves Tivdak - First Tissue Factor (TF)-Targeted Antibody Conjugate Drug (ADC) | Biopharma PEG. https://www.biochempeg.com/article/219.html.

29. Seagen Inc. *A Single Arm, Multicenter, International Trial of Tisotumab Vedotin (HuMax®-TF-ADC) in Previously Treated, Recurrent or Metastatic Cervical Cancer*. https://clinicaltrials.gov/ct2/show/NCT03438396 (2022).

30. Seagen Inc. *A Multi-center, Open-label Trial Investigating the Efficacy and Safety of Continued Treatment With Tisotumab Vedotin in Patients With Solid Tumors Known to Express Tissue Factor.* https://clinicaltrials.gov/ct2/show/NCT03245736 (2021).

31. Seagen Inc. *Dose-escalating and Cohort Expansion Safety Trial of Tissue Factor Specific Antibody Drug Conjugate Tisotumab Vedotin (HuMax®-TF-ADC) in Patients With Locally Advanced and/or Metastatic Solid Tumors Known to Express Tissue Factor*. https://clinicaltrials.gov/ct2/show/NCT02552121 (2021).

32. Seagen Inc. *First-in-human, Dose-escalating Safety Study of Tissue Factor Specific Antibody Drug Conjugate Tisotumab Vedotin (HuMax® TF ADC) in Patients With Locally Advanced and/or Metastatic Solid Tumors Known to Express Tissue Factor*. https://clinicaltrials.gov/ct2/show/NCT02001623 (2021).

33. Gambaro, K. *et al.* VGLL3 expression is associated with a tumor suppressor phenotype in epithelial ovarian cancer. *Mol. Oncol.* **7**, 513–530 (2013).

34. Yamaguchi, N. Multiple Roles of Vestigial-Like Family Members in Tumor Development. *Front. Oncol.* **10**, 1266 (2020).

35. Ren, X. Immune Microenvironment and Response in Prostate Cancer Using Large Population Cohorts. *Front. Immunol.* **12**, 16 (2021).

36. Hori, N. *et al.* Vestigial-like family member 3 (VGLL3), a cofactor for TEAD transcription factors, promotes cancer cell proliferation by activating the Hippo pathway. *J. Biol. Chem.* **295**, 8798–8807 (2020).

37. Hori, N. *et al.* Vestigial-like family member 3 stimulates cell motility by inducing high-mobility group AT-hook 2 expression in cancer cells. *J. Cell. Mol. Med.***26**, 2686–2697 (2022).

38. Zhang, L.-H. *et al.* Vestigial like family member 3 is a novel prognostic biomarker for gastric cancer. *World J. Clin. Cases* **7**, 1954–1963 (2019).

39. Zhang, L., Li, L., Mao, Y. & Hua, D. VGLL3 is a prognostic biomarker and correlated with clinical pathologic features and immune infiltrates in stomach adenocarcinoma. *Sci. Rep.* **10**, 1355 (2020).

40. Cooperberg, M. R., Broering, J. M. & Carroll, P. R. Risk Assessment for Prostate Cancer Metastasis and Mortality at the Time of Diagnosis. *JNCI J. Natl. Cancer Inst.* **101**, 878–887 (2009).

41. Kattan, M. W., Eastham, J. A., Stapleton, A. M. F., Wheeler, T. M. & Scardino, P. T. A Preoperative Nomogram for Disease Recurrence Following Radical Prostatectomy for Prostate Cancer. *JNCI J. Natl. Cancer Inst.* **90**, 766–771 (1998).

42. D'Amico, A. V. Risk-based management of prostate cancer. *N. Engl. J. Med.* **365**, 169–171 (2011).

43. Schaeffer, E. *et al.* NCCN Guidelines Insights: Prostate Cancer, Version 1.2021: Featured Updates to the NCCN Guidelines. *J. Natl. Compr. Canc. Netw.* **19**, 134–143 (2021).

44. Eastham, J. A. *et al.* Clinically Localized Prostate Cancer: AUA/ASTRO Guideline, Part I: Introduction, Risk Assessment, Staging, and Risk-Based Management. *J. Urol.* **208**, 10–18 (2022).

45. Mottet, N. *et al.* EAU-EANM-ESTRO-ESUR-SIOG Guidelines on Prostate Cancer—2020 Update. Part 1: Screening, Diagnosis, and Local Treatment with Curative Intent. *Eur. Urol.***79**, 243–262 (2021).

46. Epstein, J. I. *et al.* The 2014 International Society of Urological Pathology (ISUP) Consensus Conference on Gleason Grading of Prostatic Carcinoma: Definition of Grading Patterns and Proposal for a New Grading System. *Am. J. Surg. Pathol.* **40**, 244–252 (2016).

47. Epstein, J. I. *et al.* A Contemporary Prostate Cancer Grading System: A Validated Alternative to the Gleason Score. *Eur. Urol.* **69**, 428–435 (2016).

48. Samaratunga, H. *et al.* The prognostic significance of the 2014 International Society of Urological Pathology (ISUP) grading system for prostate cancer. *Pathology (Phila.)* **47**, 515–519 (2015).

49. Offermann, A. *et al.* Prognostic Value of the New Prostate Cancer International Society of Urological Pathology Grade Groups. *Front. Med.* **4**, 157 (2017).

50. Zelic, R. *et al.* Predicting Prostate Cancer Death with Different Pretreatment Risk Stratification Tools: A Head-to-head Comparison in a Nationwide Cohort Study. *Eur. Urol.* **77**, 180–188 (2020).