Clinical utility of human glandular kallikrein 2 within a neural network for prostate cancer detection

CARSTEN STEPHAN*,, KLAUS JUNG*, ANTONINUS SOOSAIPILLAI**, GEORGE M. YOUSEF**,†, HENNING CAMMANN‡, HELLMUTH MEYER*, CHUANLIANG XU*,§, and ELEFTHERIOS P. DIAMANDIS**,¶**

**Department of Urology and ‡Institute for Medical Biometry, University Hospital Charité, Berlin, Germany, ** Department of Pathology and Laboratory Medicine, Mount Sinai Hospital, and ¶Department of Laboratory Medicine and Pathobiology, University of Toronto, Toronto, and †Department of Laboratory Medicine, Memorial University, St. John's, Newfoundland, Canada, §Department of Urology, Changhai Hospital, Second Military Medical University, Shanghai, China*

Accepted for publication 5 April 2005

OBJECTIVE

To assess, using artificial neural networks (ANNs), human glandular kallikrein 2 (hK2), prostate-specific antigen (PSA), and percentage free/total PSA (f/tPSA), for discriminating between prostate cancer and benign prostatic hyperplasia (BPH).

MATERIAL AND METHODS

Serum samples from 475 patients with prostate cancer ($n = 347$) or BPH ($n = 128$) within the PSA range of 1–20 ng/mL were analysed for tPSA, fPSA and hK2 (research assay, Toronto, Canada). Data were analysed in the ranges of 1–4, 2–4, 4–10, and 2–20 ng/mL tPSA. Back-propagation ANN models with the variables PSA, f/tPSA, and hK2, hK2/fPSA and hK2/(f/tPSA) were constructed. The diagnostic validity was evaluated by receiver-operating characteristic (ROC) curve analysis.

RESULTS

Whereas the median concentration of hK2 was not significantly different between patients with BPH or prostate cancer in any of the tPSA ranges, the f/tPSA, hK2/fPSA and hK2/(f/tPSA), and the hK2-based ANN outputs were always significantly different between patients with prostate cancer or BPH. Using ROC curve comparison, all variables were significantly better than hK2 in all ranges. The hK2-based ANN performed better than f/tPSA except in the 4–10 ng/mL tPSA range. At 90% and 95% sensitivity, the hK2-based ANN was also significantly better than f/tPSA in the 1–4 ng/mL tPSA range. hK2/(f/tPSA) achieved equal results to the hK2-based ANN except in the range 2–20 ng/mL tPSA.

CONCLUSIONS

The hK2-based ANN improves the outcome of f/tPSA but not hK2/(f/tPSA) in almost all analysed subgroups. When comparing the results at 90% and 95% sensitivity the hK2-based ANN only performed significantly better than f/tPSA in the lowest tPSA range. Only in lower tPSA ranges do hK2-based ANNs show an advantage for further improving prostate cancer detection.

KEYWORDS

prostate cancer, PSA, human glandular kallikrein 2, artificial neural network

INTRODUCTION

Assessing PSA levels improves the early detection of prostate cancer, one of the most commonly diagnosed cancers in men [1]. Elevated PSA levels are not unique to patients with prostate cancer, as they can also occur in patients with no malignancy, e.g. BPH or chronic prostatitis [2]. Apart from the substantial overlap between prostate cancer and BPH in patients with PSA levels in the diagnostic 'grey zone' of 4–10 ng/mL, 20–30% of patients with prostate cancer have PSA concentrations of <4 ng/mL [3–5]. Recent data from a large screening study with 2950 biopsied men after 7 years follow-up with PSA concentrations <4 ng/mL had an increasing incidence of prostate cancer in the PSA ranges of 0–0.5 ng/mL (6.6% incidence), 0.6–1 (10.1%), 1.1–2 (17%), 2.1–3 (23.9%), and 3.1–4 (26.9%) [5]. Thus, men with PSA concentrations of 2–4 ng/mL have a similar incidence of prostate cancer as men with PSA concentrations of 4–10 ng/mL. The use of the percentage free/total PSA (f/tPSA) can increase the specificity for prostate cancer detection by ≈20% at PSA concentrations of 4–10 ng/mL [6]. At PSA concentrations <4 ng/mL, f/tPSA can also improve the cancer detection rate [7,8]. Other data at PSA concentrations of 2.6–4 ng/mL [9] and 2–4 ng/mL showed only a small advantage of using f/tPSA to avoid unnecessary prostate biopsies [10].

PSA (or human kallikrein 3) belongs to the extended human kallikrein family that was recently given a new nomenclature [11]. Human glandular kallikrein 2 (hK2), another member of the human kallikrein family with the highest homology to PSA (≈ 80% identity at the amino acid and DNA level) was reported to have additional value especially for the early detection of prostate cancer [12–16]. The ratio of hK2 to fPSA and the combination of f/tPSA and hK2/fPSA within the PSA concentration ranges of 2–4 and 4–10 ng/mL enhances the discrimination between patients with prostate cancer or BPH [16]. Other studies have confirmed the advantage of using hK2 and its ratios in addition to fPSA and f/tPSA, especially at low PSA concentrations for detecting prostate cancer [12,14,15]. Also, hK2 shows potential to discriminate between high- and low-grade tumours, and stage 2 and 3 tumours [17,18].

Another approach to improve prostate cancer detection rates is the use of different models of logistic regression [19,20] and artificial neural networks (ANNs), by using f/tPSA together with PSA [21–24]. However, logistic regression has a limited capacity to handle very complex data, and ANNs are better for detection, especially of nonlinear relationships among multiple variables [25].

These and other ANNs were trained with different input variables such as tPSA, f/tPSA, complexed PSA, age, race, family history, prostate volume, prostate volume indexes, findings from TRUS, or status of DRE [21–24,26–30]. As prostate volume (measured by TRUS) and status of DRE are more subjective measurements, they were not included in the present study. f/tPSA-based ANNs at low PSA levels (<4 ng/mL) have already clearly shown an increase in specificity of up to ≈50% at high sensitivity levels [22] or an increase in sensitivity at 95% specificity [24,30].

Thus, hK2 and f/tPSA-based ANNs can be used to avoid unnecessary prostate biopsies for prostate cancer detection [12,15,16,21,24,30]. The aim of the present study was: (a) to evaluate hK2 and PSA compared to f/tPSA, hK2/fPSA and hK2/(f/tPSA); (b) to develop an hk2- and f/tPSA-based ANN with the input factors tPSA, f/tPSA, hK2, hK2/fPSA, and hK2/(f/tPSA) to compare it to an ANN with only tPSA and f/tPSA (without hK2); and (c) to establish whether ANNs including hK2 at different PSA ranges can help to eliminate unnecessary biopsies, especially at lower PSA ranges.

TABLE 1 Patients investigated in the four different tPSA ranges and median values for tPSA, f/tPSA, hK2, hK2/fPSA, hK2/(f/tPSA), and hK2-ANN outputs

	tPSA, ng/mL			
Group/variable	1–4	2–4	4–10	2–20
N patients	128	83	232	430
Prostate cancer				
n	65	52	183	334
tPSA, ng/mL	2.89*	3.0	6.31	7.15*
f/tPSA, %	13.68*	14.06*	9.25*	9.34*
hK2, ng/L	136.7	154.6	225.2	228.7
hK2/fPSA	416.4*	410.8*	388.6*	366.9*
hK2/f/tPSA	10.63*	11.67*	24.78*	25.71*
hK2-ANN	0.65*	0.80*	0.84*	0.90*
BPH				
n	63	31	49	96
tPSA, ng/mL	1.93	2.97	5.79	5.2
f/tPSA, %	23.66	21.31	17.29	17.32
hK2, ng/L	125.7	160.8	232.7	207.2
hK2/fPSA	331.4	272.6	205.2	204.3
hK2/f/tPSA	5.2	6.38	13.68	12.32
hK2-ANN	0.36	0.51	0.74	0.52

$P < 0.05$ vs BPH

MATERIALS AND METHODS

The study included 475 urologically referred patients with a mean (range) age of 64 (43–86) years with prostate cancer ($n = 347$) or histologically-confirmed BPH ($n = 128$) with PSA levels of 1–20 ng/mL. All patients had a TRUS-guided sextant or octant biopsy. The histological outcome of the patients with no malignancy (BPH group) was BPH with or without chronic prostatitis, but no evidence of prostatic intraepithelial neoplasia. The criteria for prostate biopsy in patients with <4 ng/mL tPSA were either a suspicious DRE, abnormal TRUS findings, or abnormal tPSA velocity. The distributions of the patients within the different PSA ranges are shown in Table 1. Serum samples were collected between 1997 and 2001 at the University Hospital Charité Berlin, Germany, and stored at −70 °C until measurement. The tPSA and fPSA were assayed using the Immulite PSA and Immulite Free PSA assays (DPC, Los Angeles, CA, USA). hK2 was measured in Toronto (Canada) with a recently published research assay [31]. A correlation study of 21 samples and six standards with the Beckman Coulter hK2 research assay [32] indicated 55% higher median hK2 values with the Toronto research assay (Passing/Bablock: $y = 1.46x + 9.49$, $r = 0.996$).

From the measured concentrations of tPSA, fPSA and hK2, the following ratios were calculated: f/tPSA, hK2/fPSA and hK2/(f/tPSA) (equivalent to hK2 × tPSA/fPSA). Data were analysed in the tPSA ranges of 1–4 and 2–4 ng/mL, for a detailed analysis of these low tPSA ranges and in the tPSA ranges of 2–20 and 4–10 ng/mL. For all tPSA ranges, ANN models were constructed with the MATLAB Neural Network Toolbox (The Mathworks, Natick, Mass, USA). A back-propagation network was applied in which the input layer consisted of the five variables: tPSA, f/tPSA, hK2, hK2/fPSA and hK2/(f/tPSA) (termed the hK2-ANN) or the variables tPSA and f/tPSA only (termed the PSA-ANN). Various hK2-ANN models including only tPSA, f/tPSA and hK2, and including the hk2 ratios hK2/fPSA, hK2/tPSA and hK2/(f/tPSA), and with two to five hidden layers were evaluated. The best performance of the hK2-ANN was reached with the five-input layer tPSA, f/tPSA, hK2, hK2/fPSA and hK2/(f/tPSA), and no logarithmic transformation of the variables. The use of three (hK2-ANN) or two neurones (PSA-ANN) as a hidden layer was found to be optimal, and the output layer had one neurone representing the output value as the probability of prostate cancer. Every ANN model was evaluated by the 'leave-one-out' method, the logical extreme of cross validation. The data set of N patients was divided into a training set with N-1 patients, i.e. all except one, who formed the test set. Permutations were then made so that each patient in turn was used as the test set and the other N-1 patients formed the training set. For each of the N-1 training sets an ANN was computed and applied to the separated patient in the test set. In this way N tests resulted in a bias-free estimator for the error rate of the ANN. Especially in small patient groups like those with a tPSA of 2–4 ng/mL, the benefit of this validation technique is the large training set, which is an advantage over our formerly used 10-fold cross validation [23,30].

The nonparametric Kruskal–Wallis test of variance, the Mann–Whitney U-test, and the Spearman rank correlation coefficient were used to assess the results statistically. The diagnostic validity of tPSA, fPSA, hK2, the ratios f/tPSA, hK2/fPSA and hK2/(f/tPSA) and the ANN output value were evaluated by receiver-operating characteristic (ROC) curve analysis. Areas under the ROC curves (AUC) were compared using the GraphROC 2.1 for Windows [33] and MedCalc 7.2 (MedCalc

TABLE 2 Areas under ROC curves for hK2, tPSA, f/tPSA, hK2/fPSA, hK2/(f/tPSA) and the hK2-based ANN in all tPSA ranges

tPSA, ng/mL	Area under ROC curve, (SEM)					
	hK2	tPSA	f/tPSA	hK2/fPSA	hK2/(f/tPSA)	hK2-ANN
1–4	0.570 (0.051)‡	0.685 (0.047)†	0.710 (0.046)*	0.666 (0.048)	0.757 (0.043)	0.753 (0.044)
2–4	0.516 (0.068)‡	0.565 (0.066)‡	0.655 (0.060)†	0.673 (0.062)*	0.685 (0.060)	0.736 (0.055)
4–10	0.535 (0.052)‡	0.564 (0.046)‡	0.784 (0.039)	0.770 (0.037)	0.770 (0.036)	0.781 (0.035)
2–20	0.561 (0.035)‡	0.629 (0.033)‡	0.795 (0.025)‡	0.737 (0.028)‡	0.768 (0.025)†	0.859 (0.019)

**P < 0.05; † P <0.01; ‡ P <0.001 vs the ANN.*

Software, Mariakerke, Belgium). Significance was defined at $P < 0.05$.

RESULTS

For all the tPSA ranges, hK2 was not significantly different between patients with BPH or prostate cancer ($P = 0.07–0.81$, Table 1). At a tPSA of 2–4 and 4–10 ng/mL, tPSA also was not significant ($P = 0.33$ and 0.17), but in the other ranges the values differed significantly between groups (both $P < 0.001$). The ratios f/tPSA, hK2/fPSA and hK2/(f/tPSA) were, in every calculation, significantly different between patients with prostate cancer or BPH ($P < 0.001–0.019$). The ratio hK2/tPSA never reached significance ($P = 0.23–0.97$). For age, there was no significant difference between patients with prostate cancer and those with BPH at the lower tPSA ranges of 1–4 and 2–4 ng/mL ($P = 0.25$ and 0.61). However, in the two other tPSA ranges, which included more patients, the age differed significantly ($P = 0.004$ and 0.03), with a median age of 65 years for patients with BPH and 64 years for patients with prostate cancer.

ROC analyses of each tPSA range for hK2, tPSA, f/tPSA, hK2/fPSA, hK2/(f/tPSA) and for the respective ANN models in all tPSA ranges used the output value. The AUCs and the significance levels compared to the hK2-ANN output are shown in Table 2. For the AUC comparison, hK2 was always outperformed by all other variables except tPSA in the ranges of 4–10 and 2–4 ng/mL tPSA ($P = 0.24$ and 0.19). tPSA only reached similar results to f/tPSA, hK2/fPSA and hK2/(f/tPSA) in the range of 1–4 ng/mL tPSA ($P = 0.08–0.74$). However, applying the recently recommended matching procedure to avoid unequal tPSA distributions when comparing f/tPSA or other new markers to tPSA [34], the AUC for tPSA (0.50) at tPSA 1–4 ng/mL was significantly less ($P < 0.001$) than the AUCs for f/tPSA (0.69), hK2/fPSA (0.72), hK2/(f/tPSA) (0.71) or the hK2-ANN (0.74). If the comparison is done without matching each patient with prostate cancer to each with BPH with the nearest possible tPSA concentration, then the AUC for tPSA (0.685, see Table 2) was not significantly different to that for f/tPSA (0.71), but only because more patients with BPH have relatively low tPSA concentrations, giving a subsequent tPSA misdistribution. In all other ranges, tPSA already had significantly smaller AUCs than f/tPSA, both hK2 ratios and the hK2-ANN. Thus, the matching procedure was not necessary for all other comparisons. As seen in Table 2, f/tPSA was outperformed in all analysed tPSA ranges by the hK2-ANN except for the range of 4–10 ng/mL tPSA ($P = 0.88$). Comparing f/tPSA to the hK2 ratios (hK2/fPSA and hK2/(f/tPSA)), there were no significant differences ($P = 0.13–0.72$) except for the range of 2–20 ng/mL tPSA, where f/tPSA was significant compared to hK2/fPSA ($P = 0.005$). The ratio hK2/(f/tPSA) gave results equal to the hK2-ANN in the ranges of 1–4, 2–4 and 4–10 ng/mL tPSA, but not in the largest tPSA range of 2–20 ng/mL.

The PSA-ANN was equal to f/tPSA, and only better than f/tPSA in the range of 1–4 ng/mL tPSA (AUC: 0.735, $P = 0.045$). The AUCs for the PSA-ANN with tPSA and f/tPSA in the tPSA ranges of 2–4 (0.659), 4–10 (0.771) and 2–20 ng/mL (0.793) were equal to f/tPSA. In the range of 1–4 ng/mL tPSA ($P = 0.074$) and in the range of 4–10 ng/mL tPSA ($P = 0.11$) the hK2-ANN was not significantly better than the PSA-ANN. However, at 2–4 and 2–20 ng/mL tPSA, the hK2-ANN outperformed the PSA-ANN.

The specificities of hK2, tPSA, f/tPSA, hK2/fPSA, hK2/(f/tPSA) and the hK2-ANN for the given sensitivities at 95% and 90% are summarized in Table 3. It is evident that there are large differences for the significance levels compared to the ANN. In the range of 1–4 ng/mL tPSA, only hK2/(f/tPSA) performed as well as the hK2-ANN at both sensitivity limits. Conversely, there were no significant differences between all variables in the range of 2–4 ng/mL tPSA at 90% and 95% sensitivity. At these limits, f/tPSA was never significantly different to the hK2-ANN in the ranges of 2–4, 4–10 and 2–20 ng/mL tPSA. The hK2/(f/tPSA) ratio also performed as well as the hK2-ANN at these limits in the ranges of 1–4, 2–4 and 4–10 ng/mL tPSA. However, in the range of 2–20 ng/mL tPSA, the hK2-ANN outperformed all others except for the f/tPSA.

Data at 90% and 95% specificity are given for all four tPSA ranges in Table 4. However, only the low tPSA ranges of 1–4 and 2–4 ng/mL are of interest for these limits, and are consequently described below. The hK2-ANN almost always outperformed hK2 and tPSA but was never different from f/tPSA in the low tPSA ranges. There was no improvement for the hK2-ANN if compared to hK2/(f/tPSA). The hK2-ANN performed significantly better than all variables at both specificity limits in the range of 2–20 ng/mL tPSA.

Despite some advantage for the ANN models in the lower tPSA ranges compared to f/tPSA, the outcome of the ratio hK2/(f/tPSA) and the hK2-ANN for the AUC was quite similar, with a significant difference only for the large tPSA range of 2–20 ng/mL. At 90% and 95% sensitivity, f/tPSA performed as well as the hK2-ANN in all tPSA ranges between 2 and 20 ng/mL. At 90% sensitivity, hK2/(f/tPSA) also performed identically to the hK2-ANN, except in the 2–20 ng/mL tPSA range. At 90% and 95% specificity the hK2-ANN could also not improve upon f/tPSA or hK2/(f/tPSA) when analysing the lower tPSA ranges.

TABLE 3 Specificities at 90% and 95% sensitivity for hK2, tPSA, f/tPSA, hK2/fPSA, hK2/(f/tPSA) and the hK2-based ANN

tPSA, ng/mL	Sensitivity %	Specificities, % (95% CI) hK2	tPSA	f/tPSA	hK2/fPSA	hK2/(f/tPSA)	hK2-ANN
1–4	90	9.5 (4.3–18.2)‡	28.6 (19.4–39.5)†	12.7 (6.5–21.9)†	22.2 (14–32.7)*	31.7 (22.2–42.8)	42.9 (32.3–54)
	95	4.8 (1.3–12.1)†	15.9 (8.9–25.6)*	11.1 (5.4–20.1)*	7.9 (3.2–16.2)*	12.7 (6.5–21.9)	25.4 (16.7–36.1)
2–4	90	25.8 (13.6–42)	16.1 (6.6–31.3)	6.5 (1.1–19.5)	29 (16.1–45.4)	19.4 (8.8–35)	25.8 (13.6–42)
	95	6.5 (1.1–19.5)	16.1 (6.6–31.3)	6.5 (1.1–19.5)	9.7 (2.7–23.6)	9.7 (2.7–23.6)	9.7 (2.7–23.6)
4–10	90	26.5 (16.5–39)†	10.2 (4.2–20.5)†	53.1 (40.4–65.4)	6.5 (16.5–39)†	36.7 (25.3–49.5)	46.9 (34.6–59.6)
	95	8.2 (2.9–18)†	6.1 (1.7–15.4)†	22.4 (13.2–34.6)	20.4 (11.6–32.3)*	22.4 (13.2–34.6)	32.7 (21.7–45.4)
2–20	90	25 (17.9–33.4)‡	22.9 (16.1–31.2)‡	41.7 (33.2–50.6)	31.3 (23.5–40)‡	34.4 (26.4–43.2)†	51 (42.2–59.8)
	95	13.5 (8.2–20.8)‡	12.5 (7.4–19.6)†	28.1 (20.7–36.7)	16.7 (10.8–24.3)†	17.7 (11.7–25.5)†	31.3 (23.5–40)

*P *< 0.05;* † P *< 0.01;* ‡ P *< 0.001 vs the ANN.*

TABLE 4 Sensitivities at 90% and 95% specificity for hK2, tPSA, f/tPSA, hK2/fPSA, hK2/(f/tPSA) and the ANN

tPSA, ng/mL	Specificity %	Sensitivities, % (95% CI) hK2	tPSA	f/tPSA	hK2/fPSA	hK2/(f/tPSA)	hK2-ANN
1–4	90	15.4 (8.6–24.9)†	21.5 (13.6–31.7)†	38.5 (28.4–49.5)	30.8 (21.5–41.6)	38.5 (28.4–49.5)	38.5 (28.4–49.5)
	95	7.7 (3.1–15.7)†	15.4 (8.6–24.9)*	27.7 (18.8–38.3)	21.5 (13.6–31.7)	33.8 (24.2–44.8)	27.7 (18.8–38.3)
2–4	90	7.7 (2.7–17)†	19.2 (10.9–30.6)†	30.8 (20.4–43)	25 (15.5–36.9)†	36.5 (25.4–48.9)	42.3 (30.7–54.7)
	95	5.8 (1.6–14.6)†	11.5 (5.2–21.8)*	25 (15.5–36.9)	7.7 (2.7–17)*	13.5 (6.6–24)	21.2 (12.4–32.7)
4–10	90	7.1 (4.3–11.1)‡	15.3 (11.1–20.4)‡	33.3 (27.6–39.5)†	48.1 (41.8–54.4)	49.2 (42.9–55.5)	50.3 (44–56.6)
	95	3.3 (1.5–6.5)‡	11.5 (7.9–16.2)‡	21.3 (16.5–26.9)†	34.4 (28.6–40.7)†	32.8 (27.1–39)†	43.7 (37.5–50.1)
2–20	90	12 (9.2–15.3)‡	13.2 (10.3–16.6)‡	42.8 (38.3–47.5)‡	45.5 (40.9–50.2)‡	47.3 (43.7–52)‡	68.6 (64.1–72.8)
	95	6.6 (4.5–9.3)‡	9.9 (7.3–13)‡	30.5 (26.4–35)‡	21 (17.4–25)‡	31.4 (27.3–35.9)‡	57.8 (53.1–62.3)

*P *< 0.05;* † P *< 0.01;* ‡ P *< 0.001 vs the ANN.*

DISCUSSION

The use of f/tPSA improves the specificity for tPSA, especially in the 4–10 ng/mL range and at tPSA concentrations <4 ng/mL [6–9]. In the present study in all patient groups, calculating f/tPSA provided a substantial improvement in the AUCs compared to tPSA except for the range of 1–4 ng/mL tPSA. However, when applying the matching procedure [34] in the range of 1–4 ng/mL tPSA, the AUC for tPSA is 0.50 instead of 0.685, and subsequently there is also a significant benefit to using f/tPSA (AUC 0.69 instead of 0.71). Similar data have been partly published elsewhere [23] for the performance of f/tPSA.

There was a further improvement when hK2 and its ratios were added to tPSA and f/tPSA [12,15,16]. In the present study, hK2 alone could not enhance the specificity of tPSA or f/tPSA as a single variable. This is in agreement with data on 937 patients reported by Partin *et al.* [16], who found no difference in hK2 levels between patients with prostate cancer or BPH, although other studies on fewer patients have reported a difference [12,15]. The ratio hK2/fPSA was significantly different between patients with prostate cancer or BPH, but did not improve the performance compared to f/tPSA (ROC analysis). The ratio hK2/(f/tPSA) performed better than hK2/fPSA (a larger AUC in three of four comparisons) and was only once outperformed by the hK2-ANN for the ROC analysis in the range of 2–20 ng/mL tPSA. This ratio of hK2/(f/tPSA) improved the discrimination between prostate cancer and BPH in another study but was no different from f/tPSA [12]. We found the same behaviour, with no differences between hK2/(f/tPSA) and f/tPSA.

Nevertheless, a key result of the present study is the better performance of the hK2-ANN than f/tPSA. In three of four ROC comparisons, the hK2-ANN performed significantly better than f/tPSA. Only in the tPSA range of 4–10 ng/mL was there no difference between f/tPSA and the hK2-ANN. Thus, especially for patients with a tPSA of <4 ng/mL, the use of hK2-ANNs has a significant advantage over f/tPSA. These findings might become more important, as recent data indicate a high prevalence (23.9–26.9%) of prostate cancer in men with 2–4 ng/mL tPSA [5]. The additional use of hK2 in combination with an ANN therefore seems promising to improve patient selection for prostate biopsy, especially for a tPSA of <4 ng/mL. Even if the advantage of the hK2-ANN to the ratio hK2/(f/tPSA) was only marginal, only these ANN models reached significance for f/tPSA, but not hK2/(f/tPSA). Notably, the greatest advantage of the hK2-ANN with significance to all other variables (Table 2) was achieved using the largest patient group at the range of 2–20 ng/mL tPSA ($n = 430$). Further analysis with more

patients within the subgroups might also improve the outcome of a trained hK2-ANN, but the wide PSA range might also contribute to this good result. To our knowledge, a combined analysis of hK2 measurements within an ANN has not been reported to date. The equal performance of the PSA-ANN with only tPSA and f/tPSA in the tPSA ranges of 1–4 and 4–10 ng/mL to the hK2-ANN limits the advantage of the hK2-based ANN.

There are some other limitations of the hK2-ANN results; looking at the clinically important limits of 90% and 95% sensitivity, the advantages of the ANN models decrease. In all tPSA ranges except 1–4 ng/mL tPSA, the hK2-ANN had no advantage over f/tPSA (Table 3). For lower tPSA concentrations, the 90% or 95% specificity limits became more important. Two ANN studies [23,24] have already emphasized the clinical usefulness of high specificity limits to avoid many false-positive tests. Again, in the present study the ANN had no advantage over f/tPSA in the two tPSA ranges up to 4 ng/mL. These findings show the limitations of the serum variable hK2 even if a powerful ANN based on this and other variables is generated. The distribution between prostate cancer and BPH, where the larger proportion is presented by prostate cancer (73% of all patients) also influences ROC analyses. In the range of 1–4 ng/mL tPSA, the incidence of prostate cancer and BPH is almost equal, but prostate cancer is over-represented in higher tPSA ranges, and this influences the positive and negative predictive values but not the sensitivity or specificity or AUC values. In the present study we therefore calculated the latter variables (Tables 2–4) but not the predictive values. However, in comparison to screening populations with 25–30% of patients with prostate cancer, the 70–75% of patients with prostate cancer in the present study represents the real proportions in our clinic as a cancer centre. Therefore, the ANN results of this investigation might not be transferable to other (screening) populations but might be useable in our future studies.

There are no commercially available hK2 assays, and there is no universally accepted calibrator for hK2 as there is for PSA. Three hK2 assays have been used for most hK2 testing to date: the assay in the present study [31]; the Beckman Coulter total hK2 assay [35]; and the Turku hK2 assay [36]. The present assay measured about 1.5 times higher hK2 concentrations than the Beckman Coulter assay. The latter assay and the Turku hK2 assay were recently compared by Haese *et al.* [37] and Blijenberg *et al.* [38], and the differences between them could not be attributed to calibration differences. As these three hK2 assays are not commercially available, none have been fully developed to reduce variability in the way that commercially available PSA and fPSA assays have been. Additionally, the recognition of free and complexed hK2, where a newly developed assay was recently described [39], is not known from the research assay used in the present study, which might also cause differences.

Although almost all studies to date have indicated promising results with hK2, there are general concerns about hK2 as a new or even as an additional marker for diagnosing prostate cancer. Promising results with hK2 have been published since 1998, but so far no hK2 assay has been used as a routine clinical assay. The limited advantage of hK2 compared to f/tPSA might be a reason for this; if hK2 itself cannot improve the discrimination between prostate cancer and BPH, the hK2-based ANN can only marginally improve this discrimination. This has been shown only in the tPSA ranges of 2–4 and 2–20 ng/mL, comparing it to the PSA-ANN.

Stamey *et al.* [40] could see no rationale as to why hK2, with its 80% homology to PSA, should independently improve prostate cancer detection beside PSA and f/tPSA. Nevertheless hK2 is a very different protein biochemically and physiologically, and is found primarily in the free form, unlike PSA which is found primarily in the complexed form [41]. As hK2 appears to have value in particular patients and in particular PSA ranges when combined with tPSA and fPSA assay data, perhaps its full value will not be realized until it is incorporated into a commercially available standardized multiplexed assay with these and other analytes [42].

In conclusion, the present data show that the hK2-ANN improves the outcome of f/tPSA in almost all analysed subgroups based on the AUC. The best performance of the hK2-ANN was seen with all patients and at low tPSA concentrations (<4 ng/mL). The advantage of the hK2-ANN decreased if the hK2/(f/tPSA) ratio or the PSA-ANN was compared to the hK2-ANN. The results at 90% and 95% sensitivity and specificity showed no improvement in the hK2-ANN performance compared to f/tPSA or hK2/(f/tPSA) except in the range of 1–4 ng/mL tPSA. However, the results should be interpreted with caution until larger patient groups, especially including more patients with no cancer, have been analysed. In patients with low tPSA values the use of hK2 in combination with an ANN might improve the discrimination between prostate cancer and BPH.

ACKNOWLEDGEMENTS

This work was supported by the Mildred-Scheel-Foundation (Grant 70–3295-ST1), and by the SONNENFELD-Stiftung. We gratefully thank Paul E.C. Sibley for the helpful corrections and Judith A. Finlay for comments on the manuscript and for correlation studies between our hK2 assay and the Beckman Coulter hK2 assay.

CONFLICT OF INTEREST

None declared.

REFERENCES

1 **Polascik TJ, Oesterling JE, Partin AW.** Prostate specific antigen: a decade of discovery – what we have learned and where we are going. *J Urol* 1999; **162**: 293–306

2 **Stephan C, Jung K, Lein M, Sinha P, Schnorr D, Loening SA.** Molecular forms of prostate-specific antigen and human kallikrein 2 as promising tools for early diagnosis of prostate cancer. *Cancer Epidemiol Biomarkers Prev* 2000; **9**: 1133–47

3 **Catalona WJ, Smith DS, Ratliff TL *et al.*** Measurement of prostate-specific antigen in serum as a screening test for prostate cancer. *N Engl J Med* 1991; **324**: 1156–61

4 **Gann PH, Hennekens CH, Stampfer MJ.** A prospective evaluation of plasma prostate-specific antigen for detection of prostatic cancer. *JAMA* 1995; **273**: 289–94

5 **Thompson IM, Pauler DK, Goodman PJ *et al.*** Prevalence of prostate cancer among men with a prostate-specific antigen level < or =4.0 ng per milliliter. *N Engl J Med* 2004; **350**: 2239–46

6 **Catalona WJ, Partin AW, Slawin KM *et al.*** Use of the percentage of free prostate-specific antigen to enhance differentiation of prostate cancer from benign prostatic disease: a prospective

multicenter clinical trial. *JAMA* 1998; **279**: 1542–7
7 **Catalona WJ, Partin AW, Finlay JA *et al.*** Use of percentage of free prostate-specific antigen to identify men at high risk of prostate cancer when PSA levels are 2.51–4 ng/mL and digital rectal examination is not suspicious for prostate cancer: an alternative model. *Urology* 1999; **54**: 220–4
8 **Jung K, Stephan C, Elgeti U *et al.*** Molecular forms of prostate-specific antigen in serum with concentrations of total prostate-specific antigen <4 μg/L: are they useful tools for early detection and screening of prostate cancer? *Int J Cancer* 2001; **93**: 759–65
9 **Roehl KA, Antenor JA, Catalona WJ.** Robustness of free prostate specific antigen measurements to reduce unnecessary biopsies in the 2.6–4.0 ng/mL range. *J Urol* 2002; **168**: 922–5
10 **Haese A, Dworschack RT, Partin AW.** Percent free prostate specific antigen in the total prostate specific antigen 2–4 ng/mL range does not substantially increase the number of biopsies needed to detect clinically significant prostate cancer compared to the 4–10 ng/mL range. *J Urol* 2002; **168**: 504–8
11 **Diamandis EP, Yousef GM, Clements J *et al.*** New nomenclature for the human tissue kallikrein gene family. *Clin Chem* 2000; **46**: 1855–8
12 **Becker C, Piironen T, Pettersson K *et al.*** Discrimination of men with prostate cancer from those with benign disease by measurements of human glandular kallikrein 2 (HK2) in serum. *J Urol* 2000; **163**: 311–6
13 **Kwiatkowski MK, Recker F, Piironen T *et al.*** In prostatism patients the ratio of human glandular kallikrein to free PSA improves the discrimination between prostate cancer and benign hyperplasia within the diagnostic 'gray zone' of total PSA 4–10 ng/mL. *Urology* 1998; **52**: 360–5
14 **Magklara A, Scorilas A, Catalona WJ, Diamandis EP.** The combination of human glandular kallikrein and free prostate-specific antigen (PSA) enhances discrimination between prostate cancer and benign prostatic hyperplasia in patients with moderately increased total PSA. *Clin Chem* 1999; **45**: 1960–6
15 **Nam RK, Diamandis EP, Toi A *et al.*** Serum human glandular kallikrein-2 protease levels predict the presence of prostate cancer among men with elevated prostate-specific antigen. *J Clin Oncol* 2000; **18**: 1036–42
16 **Partin AW, Catalona WJ, Finlay JA *et al.*** Use of human glandular kallikrein 2 for the detection of prostate cancer: preliminary analysis. *Urology* 1999; **54**: 839–45
17 **Haese A, Graefen M, Steuber T *et al.*** Human glandular kallikrein 2 levels in serum for discrimination of pathologically organ-confined from locally-advanced prostate cancer in total PSA-levels below 10 ng/mL. *Prostate* 2001; **49**: 101–9
18 **Recker F, Kwiatkowski MK, Piironen T *et al.*** Human glandular kallikrein as a tool to improve discrimination of poorly differentiated and non-organ-confined prostate cancer compared with prostate-specific antigen. *Urology* 2000; **55**: 481–5
19 **Carlson GD, Calvanese CB, Partin AW.** An algorithm combining age, total prostate-specific antigen (PSA), and percent free PSA to predict prostate cancer: results on 4298 cases. *Urology* 1998; **52**: 455–61
20 **Virtanen A, Gomari M, Kranse R, Stenman UH.** Estimation of prostate cancer probability by logistic regression: free and total prostate-specific antigen, digital rectal examination, and heredity are significant variables. *Clin Chem* 1999; **45**: 987–94
21 **Finne P, Finne R, Auvinen A *et al.*** Predicting the outcome of prostate biopsy in screen-positive men by a multilayer perceptron network. *Urology* 2000; **56**: 418–22
22 **Babaian RJ, Fritsche H, Ayala A *et al.*** Performance of a neural network in detecting prostate cancer in the prostate-specific antigen reflex range of 2.5–4.0 ng/mL. *Urology* 2000; **56**: 1000–6
23 **Stephan C, Jung K, Cammann H *et al.*** An artificial neural network considerably improves the diagnostic power of percent free prostate-specific antigen in prostate cancer diagnosis: results of a 5-year investigation. *Int J Cancer* 2002; **99**: 466–73
24 **Djavan B, Remzi M, Zlotta A, Seitz C, Snow P, Marberger M.** Novel artificial neural network for early detection of prostate cancer. *J Clin Oncol* 2002; **20**: 921–9
25 **Partin AW, Murphy GP, Brawer MK.** Report on Prostate Cancer Tumor Marker Workshop 1999. *Cancer* 2000; **88**: 955–63
26 **Finne P, Finne R, Bangma C *et al.*** Algorithms based on prostate-specific antigen (PSA), free PSA, digital rectal examination and prostate volume reduce false-positive PSA results in prostate cancer screening. *Int J Cancer* 2004; **111**: 310–5
27 **Horninger W, Bartsch G, Snow PB, Brandt JM, Partin AW.** The problem of cutoff levels in a screened population: appropriateness of informing screenees about their risk of having prostate carcinoma. *Cancer* 2001; **91**: 1667–72
28 **Kalra P, Togami J, Bansal BSG *et al.*** A neurocomputational model for prostate carcinoma detection. *Cancer* 2003; **98**: 1849–54
29 **Remzi M, Anagnostou T, Ravery V *et al.*** An artificial neural network to predict the outcome of repeat prostate biopsies. *Urology* 2003; **62**: 456–60
30 **Stephan C, Cammann H, Semjonow A *et al.*** Multicenter evaluation of an artificial neural network to increase prostate cancer detection rate and reduce unnecessary biopsies. *Clin Chem* 2002; **48**: 1279–87
31 **Black MH, Magklara A, Obiezu CV, Melegos DN, Diamandis EP.** Development of an ultrasensitive immunoassay for human glandular kallikrein with no cross-reactivity from prostate-specific antigen. *Clin Chem* 1999; **45**: 790–9
32 **Finlay JA, Day JR, Evans CL *et al.*** Development of a dual monoclonal antibody immunoassay for total human kallikrein 2. *Clin Chem* 2001; **47**: 1218–24
33 **Kairisto V, Poola A.** Software for illustrative presentation of basic clinical characteristics of laboratory tests – GraphROC for Windows. *Scand J Clin Lab Invest Suppl* 1995; **222**: 43–60
34 **Jung K, Stephan C, Lein M *et al.*** Receiver-operating characteristic as a tool for evaluating the diagnostic performance of prostate-specific antigen and its molecular forms-What has to be considered? *Prostate* 2001; **46**: 307–10
35 **Klee GG, Goodmanson MK, Jacobsen SJ *et al.*** Highly sensitive automated chemiluminometric assay for measuring free human glandular kallikrein-2. *Clin Chem* 1999; **45**: 800–6
36 **Becker C, Piironen T, Kiviniemi J, Lilja H, Pettersson K.** Sensitive and specific immunodetection of human glandular

kallikrein 2 in serum. *Clin Chem* 2000; **46**: 198–206
37 **Haese A, Vaisanen V, Finlay JA *et al.*** Standardization of two immunoassays for human glandular kallikrein 2. *Clin Chem* 2003; **49**: 601–10
38 **Blijenberg BG, Wildhagen MF, Bangma CH, Finlay JA, Vaisanen V, Schroder FH.** Comparison of two assays for human kallikrein 2. *Clin Chem* 2003; **49**: 243–7
39 **Vaisanen V, Eriksson S, Ivaska KK, Lilja H, Nurmi M, Pettersson K.** Development of sensitive immunoassays for free and total human glandular kallikrein 2. *Clin Chem* 2004; **50**: 1607–17
40 **Stamey TA, Warrington JA, Caldwell MC *et al.*** Molecular genetic profiling of Gleason grade 4/5 prostate cancers compared to benign prostatic hyperplasia. *J Urol* 2001; **166**: 2171–7
41 **Rittenhouse HG, Finlay JA, Mikolajczyk SD, Partin AW.** Human Kallikrein 2 (hK2) and prostate-specific antigen (PSA): two closely related, but distinct, kallikreins in the prostate. *Crit Rev Clin Lab Sci* 1998; **35**: 275–368
42 **Finlay JA, Mikolajczyk SD, Pribyl TM *et al.*** Prostate Cancer Markers: From Discovery to the Clinic. In Nakamura RM, Grody WW, Wu JT, Nagle RB, eds, *Cancer Diagnostics Current and Future Trend*, Chapt 7. Totowa: Humana Press, 2004: 85–128

Correspondence to: Dr Carsten Stephan, Department of Urology, University Hospital Charité Berlin, CCM, Schumannstrasse 20/21, D-10098 Berlin, Germany.
e-mail: carsten.stephan@charite.de

Abbreviations: **hK2**, human glandular kallikrein 2; **ANN**, artificial neural network; **f/tPSA**, percentage free/total PSA; **ROC**, receiver-operating characteristic; **AUC**, area under curve.