External quality assessment of tumour marker analysis: state of the art and consequences for estimating diagnostic sensitivity and specificity

Ringversuche für Tumormarker: eine Istanalyse mit Folgerung für die Berechnung der diagnostischen Sensitivität und Spezifität

Abstract

This review shows the current analytical quality for the following analytes used as tumour markers in the external quality assessment (EQA)-programmes of Instand e.V., a national EQA-organiser in Germany: Corticotropin (ACTH), growth hormone (GH, hGH), prolactin (PRL), chorionic gonadotropin (CG, hCG), calcitonin (CT, hCT), thyroglobulin (Tg), carcinoembryonic antigen (CEA), CA-Antigens 125, 72-4, 15-3 and 19-9, alpha foetoprotein (AFP) and prostate-specific antigen (PSA).

The results from the participants show a large variation in the precision of the methods used as well as in the comparability of results between methods for the same analyte. In general, the hormones used as tumour markers show better performance than the "CA-markers", which are often inadequately standardised and defined. In the case of one CA-marker (CA 72-4/TAG 72-4), the differences between the lowest kit median concentration and highest kit median concentration for one sample pair were 440% and 580%. The corresponding figures for ACTH were 123% and 156% and for CEA 180% and 184%.

The classical tumour markers such as carcinoembryonic antigen (CEA) and alpha foetoprotein (AFP) performed markedly better than the CA-markers and PSA with regards to both inter- and intra-method comparability.

The inter-laboratory precision for a given kit and marker was acceptable in many cases.

The results show that only results from the same kit/method for each tumour marker can be used for cumulative or time-dependent comparison of results - for example pre-operative and post-operative follow up. In the case of prostate specific antigen (PSA), the kits used for free and total PSA must come from the same producer, if the generally accepted ratios are to have any diagnostic value.

The need for kit- and laboratory-specific reference ranges and cut-off values for setting diagnostic specificity and sensitivity is highlighted from the EQA-results. The situation for inter-method comparability for the CA-Markers has not improved over the past decade.

With the exception of calcitonin for detecting medullary thyroid carcinoma, chorionic gonadotropin in germ-cell tumours in men and thyroglobulin after total thyroidectomy, none of the remaining analytes appear to be suitable for screening purposes.

Keywords: external quality assessment (EQA), tumour markers, precision, accuracy, standardisation, method comparison, diagnostic specificity, diagnostic sensitivity, immunoassay

Hans Reinauer[1]
William Graham Wood[1]

1 Instand Reference Laboratory, Instand e.V., Düsseldorf, Germany

Zusammenfassung

Diese Übersicht stellt die analytische Qualität von einigen Tumormarkern in den Ringversuchen dar. Es handelt sich um die Analyte: Corticotropin

(ACTH), Wachstumshormon (GH, hGH), Prolaktin (PRL), Choriongonado-
tropin (hCG und CG), Calcitonin (CT), Thyreoglobulin (Tg), Carcinoembryo-
nales Antigen (CEA), CA-Antigene 125, 72-4, 15-3 und 19-9, Alpha-Fe-
toprotein (AFP) und Prostata-spezifisches Antigen (PSA). Die Analysen-
ergebnisse der Ringversuchsteilnehmer weisen eine relativ große Un-
präzision und mangelhafte Vergleichbarkeit bei demselben Analyten
auf. Lediglich bei den Hormonen, soweit sie als Tumormarker verwendet
werden, sind die Analysenergebnisse besser vergleichbar als diejenigen
der anderen Protein-Tumormarker, bei denen offenbar eine durchgrei-
fende Standardisierung noch aussteht. Als Beispiel sei genannt der
Analyt CA 72-4 (TAG 72-4), bei dem die Unterschiede der Medianwerte
der verschiedenen Analyseverfahren zum Teil 440% - 580% betragen.
Für ACTH waren die vergleichbaren Zahlen 123% und 156%, für CEA
180% - 184%. Die klassischen Tumormarker wie AFP und CEA erzielten
bessere Ergebnisse als die CA-Marker und PSA.

Die Ergebnisse der Studie zeigen, dass für die Langzeitüberwachung
von Patienten nur das gleiche Analysesystem bzw. der gleiche Testkit
eingesetzt werden darf. Für Prostata-spezifisches Antigen sollten die
Reagenzien zur Bestimmung von gesamt- und freiem PSA vom selben
Hersteller bezogen werden und auf äquimolarer Basis funktionieren.
Jedes Labor sollte seine eigenen Referenzbereiche und cut-off Werte
für jeden Tumormarker selbst erstellen oder vom Hersteller angeben
lassen. Die Übernahme von cut-off Werten aus Lehrbüchern für alle
Analysensysteme ist nicht angezeigt. Bedauerlicherweise hat sich die
Vergleichbarkeit zwischen den Reagenziensystemen für die CA-Marker
von verschiedenen Herstellern über die letzten Jahre nicht wesentlich
verbessert.

Mit Ausnahme von Calcitonin zur Diagnose von medullären Schilddrü-
senkarzinomen, Choriongonadotropin für Keimzelltumoren bei Männern
und Thyreoglobulin nach totaler Thyreoidektomie ist keiner der aufge-
führten Marker für das Tumor-Screening geeignet.

Introduction

The concept of using markers for the monitoring of tumour
growth and efficiency of therapeutic intervention is not
new. The first attempts at monitoring tumour growth
and/or activity were with monoclonal immunoglobulino-
pathies using electrophoretic techniques and endocrine
tumours using hormones as markers. The chemical ana-
lysis of hapten hormones such as adrenaline, noradren-
aline, serotonin in urine or plasma and their metabolites,
homovanillic acid, vanillylmandelic acid and 5-hydroxyin-
dole acetic acid in urine were used as differential markers
for suprarenal neoplasia (neuroblastoma, phaeochromo-
cytoma, carcinoid) [1], [2], [3], [4]. Tartrate-inhibited acid
phosphatase (ti-AcP) was used as an indicator of benign
or malignant prostatic disorders [5].

The introduction of immunoassays allowed the measure-
ment of peptide hormones, at first mainly for pituitary
disorders (corticotropin (ACTH), growth hormone (GH,
hGH), chorionic gonadotropin (CG, hCG), prolactin (PRL),
thyrotropin (TSH)), steroid hormones (cortisol, aldoster-
one) and thyroid hormones (thyroxine (T_4) and triiodo-
thyronine (T_3)) [6], [7], [8], [9], [10], [11], [12].

The search for more specific markers for non-endocrine
tumours led to the development of carcinoembryonic
antigen (CEA), alpha foetoprotein (AFP) and the "CA-im-
munoassays", originally from Centocor. The latter include
CA 19-9 (gastrointestinal tumours), CA 125 (ovarian
cancer), CA 15-3 (breast cancer) and CA 72-4 (TAG 72-4,
gastric tumours) [13], [14], [15], [16], [17], [18].

The use of prostate-specific antigen (PSA) as a more
specific tumour marker became interesting after the
discovery that PSA was present in at least three forms in
serum (free and complexed with either α_1-antichymotryp-
sin (α_1-ACT) and α_2-macroglobulin (α_2-MG), the latter of
which is not detectable by immunoassay, as the PSA is
surrounded by the α_2-MG and is therefore "invisible" to
antibodies). Assays were developed for total, free and
complexed PSA [19], [20], [21]. The ratio of "free" PSA to
complexed PSA (= PSA-α_1-ACT) has been propagated as
being able to differentiate between benign (hyperplasia)
and malignant (neoplastic) prostatic disease [22], [23].

With the advent of the polymerase chain reaction (PCR)
- methods, the accent of diagnosis is shifting toward the
use of genetic markers. Examples here are the BRCA-1
and BRCA-2 genes which are associated with breast and
ovarial cancer [24], [25], [26].

This publication deals with the review of tumour-marker
external quality assessment (EQA) - schemes offered by
Instand e.V. with a critical appraisal of performance and
method comparability and their influence on the estima-
tion of the diagnostic sensitivity and specificity.

The large variation in kit-performance, as far as numerical
values are concerned, raises the question as to whether

Table 1: 1-39 ACTH - EQAS results - arranged according to methods

ACTH concentrations in ng/l

Method	N	Sample 1				Sample 2			
		Median	Mean	CV %	SR %	Median	Mean	CV %	SR %
1a	25	263	274	18.4	92	474	498	15.0	84
2b/c	25	227	223	22.5	96	370	357	21.0	80
3a/c	4	278	275	18.4	100	579	556	10.8	100
99z	4	280	269	6.94	75	456	417	16.3	75

Method	N	Sample 3			
		Median	Mean	CV %	SR %
1a	42	<10	<10		100
2b/c	33	<10	<10		100
3a/c	4	1.38	1.29	19.0	100
99z	2	<10	<10		100

Key:

Method: Kits are ordered numerically – for example: Kit 1 is always the same producer. The suffix is the label used: a – enzyme (irrespective of substrate); b – chemiluminescence; c – radioisotope; d – fluorescence; z – mixture (only used in conjunction with the "remainder/rest" group 99).

N – number of participants using the kit in column 1.

CV – coefficient of variation in percent

SR – success rate in percent

Target values set by selected laboratories using immunometric assays:
Sample 1 – 278 ng/l; Sample 2 – 428 ng/l; Sample 3 – <10 ng/l
Limits set for successful participation – target value ±60% (Sample 3 – <10 ng/l)

the "cut-off" values generally accepted but not validated for each method/kit, are of any use in terms of setting limits for diagnostic sensitivity and specificity.

Materials and methods

The data used in this publication was obtained from the EQA schemes run by Instand e.V. during 2003 and 2004. Two samples - lyophilised processed serum for the proteohormones, thyroglobulin and α-foetoprotein; liquid processed serum for the CA-markers, carcinoembryonic antigen and prostate-specific antigen - containig different analyte concentrations were sent for analysis with each survey.

Participants were requested to return the analysis data together with details concerning the measuring device and method/kit used.

Data was used from EQA-schemes for the following analytes: calcitonin (CT, hCT), thyroglobulin, α-foetoprotein (AFP), chorionic gonadotropin, prolactin, corticotropin, prostate-specific antigen (free, total and complexed: f-PSA; t-PSA; c-PSA), carcinoembryonic antigen (CEA), CA-125, CA 19-9, CA 72-4 and CA 15-3.

Statistics

Any statistical comparisons used were based on the calculation of the measure of central tendency (mean, median) and dispersion (standard deviation and coefficient of variation (normally distributed data)) defined confidence intervals as centiles (non-Gaussian data distribution)

and mainly included box and whisker plots and data distribution curves for graphic illustration of group-results. Statistical comparisons between methods for a given analyte were not made because the study was designed to present data in a visual way, rather than to compare the inter-method performance.

Results

The results are given separately for each analyte. Samples 1 and 2 were the same pair in Tables 1-3. Sample 1 was pooled adult male serum, Sample 2 pooled adult female serum. Both samples were spiked and filtered (0.2 µm pore size).

Generally accepted cut-off levels for the "CA-Markers", CEA and total PSA have been given in the text [27], (see also [28]). These "decision-limits" have usually been made many years ago for assay-designs (e.g. competitive radioimmunoassay) no longer existing. They have not been validated for immunometric methods and should be interpreted with care.

There are no such "dogmatic-limits" for the other analytes used as tumour markers.

a. Corticotropin (1-39 ACTH)

The problems with 1-39 ACTH are due more to stability of the analyte than to analytical problems, although some kits react more "allergically" than others.

Table 1 shows the results from three samples sent in different EQA-surveys in 2004.

Table 2: Growth hormone - EQAS results - arranged according to methods
Growth hormone concentrations in µg/l

Method	N	Sample 1				Sample 2			
		Median	Mean	CV %	SR %	Median	Mean	CV %	SR %
99z-2	6	0.602	0.586	19.6	100	0.579	0.541	20.6	100
99z-1	44	0.980	0.982	12.4	98	0.975	0.960	16.0	98

Method	N	Sample 3			
		Median	Mean	CV %	SR %
99z-2	6	1.07	1.01	13.8	100
99z-1	57	1.60	1.57	14.5	98

Key: as in Table 1 except for:
99z-2 Kits using the 2nd international standard (IS) 2000 (NIBSC 98/574 1.95 mg protein/ampoule - recombinant)
99z-1 Kits using the 1st IS 1987 (NIBSC 80/505 4.4 IU/ampoule – human pituitary)

The target values were the median values of all participants in the group. The dispersion allowed was the median ±45%.

Table 3: Prolactin - EQAS results - arranged according to methods
Prolactin concentrations in µg/l

Method	N	Sample 1				Sample 2			
		Median	Mean	CV %	SR %	Median	Mean	CV %	SR %
4a	12	10.8	10.7	11.5	100	10.6	10.4	11.4	100
5b/6a	36	12.2	12.1	10.2	100	12.2	12.1	10.9	100
1a	21	8.64	8.79	13.0	95	8.40	8.60	11.9	90
7b	31	7.60	7.73	6.10	100	7.60	7.54	9.70	100
99z	16	8.75	8.87	17.5	100	8.72	8.75	16.2	100

Method	N	Sample 3			
		Median	Mean	CV %	SR %
4a	14	7.20	7.04	14.4	93
5b/6a	40	8.02	8.19	12.5	100
1a	24	5.41	5.49	10.8	100
7b	37	5.36	5.34	8.40	100
99z	17	5.01	5.01	20.4	94

Key: as in Table 1 except:
Target values – group median
Dispersion allowed – target value ±45% for all groups except 99z (target value ±60%)

Reference Standard – 3rd IS 1988 (NIBSC 84/500)

b. Growth Hormone (GH, hGH)

At present, kits are using both the 1st IS, which is an extract from human pitutitaries, as well as the recombinant 22 kD 2nd IS. The comparison between kits is so good that all kits using the same reference standard can be grouped together. The values obtained using the recombinant material are substantially lower than those from the extracted human pituitaries.
Table 2 shows the results for growth hormone assays from three samples sent in different EQA-surveys (EQAS) in 2004.

c. Prolactin (PRL)

Prolactin kits were calibrated exclusively with the 3rd international standard. Table 3 shows the results for prolactin results from three samples distributed in 2004.

d. Chorionic Gonadotropin (CG, hCG)

The kits were divided into those measuring only holo-hCG, holo-hCG + free β-chain and only free β-chain. Results were given as method specific in the first two cases and as a single group in the latter case.
The calibration of methods for holo-hCG was with the 3rd IS (NIBSC 75/537), for free β-chain with the 1st IRP for immunoassay (NIBSC 75/551). Accurate calibration of kits which determine both holo-hCG and free β-chain is impossible, as there are two unknown variables. Both reference preparations are given in IU/ampoule, but 1 IU holo-hCG is not equal to 1 IU free β-chain. The specificity of the antibodies for holo-hCG is well documented in Table 4. The cross-reactivity of the free β-chain in the "mixed" assays can be seen by comparing the values for samples 1 and 2 in Tables 4 and 5 . The results from three EQA samples sent in 2004 are shown in Tables 4-6 .

Table 4: Holo-hCG - EQAS results - arranged according to methods
hCG concentrations in IU/l (NIBSC 75/537)

Method	N	Sample 1				Sample 2			
		Median	Mean	CV %	SR %	Median	Mean	CV %	SR %
5b/8a	22	2.04	2.08	14.7	95	1.09	1.22	15.9	100
6a	18	<2.00	<2.00		100	<2.00	<2.00		100
99z	9	2.36	2.26	16.4	89	1.34	1.20	10.5	100

Method	N	Sample 3			
		Median	Mean	CV %	SR %
5b	13	660	655	6.58	86
7a	3	840	839	3.40	100
6a	11	729	751	8.75	100
1a	5	660	691	14.3	100
99z	7	707	750	16.6	100

Key: as in Table 1 except:
Target values – group median for Sample 3
Dispersion allowed – target value ±38% for Sample 3 all groups (Current Guideline of the German Medical Council [45]).

For Samples 1 and 2 the target value was ≤2 IU/l – calibrated with 3rd IS 1986 for chorionic gonadotropin (NIBSC 75/537) – 650 IU per ampoule.

Table 5: Holo-hCG + free β-chain - EQAS results - arranged according to methods

Method	N	Sample 1				Sample 2			
		Median	Mean	CV %	SR %	Median	Mean	CV %	SR %
4a	33	252	253	5.78	93	131	132	8.91	100
1a	13	274	278	7.42	100	143	142	6.89	100
5b	25	177	178	9.10	96	89.0	90.2	8.46	96
7b/9b	27	122	122	12.3	96	60.1	59.1	12.7	96
10a/11a/3d	6	260	239	21.2	100	135	123	19.5	100
99z	13	273	263	21.5	80	143	138	22.3	80

Method	N	Sample 3			
		Median	Mean	CV %	SR %
4a	20	870	865	7.30	100
1a/3d/10a/11a	16	704	709	9.12	100
5b	17	671	673	8.91	100
7b/12a	6	847	805	14.7	100
99z	4	743	756	22.2	100

Key: as in Table 1 except:
Target values – group median for all samples
Dispersion – ±38% (Current Guideline of the German Medical Council [45]
hCG concentrations in "IU/l" (theoretically not possible – equation with 2 unknowns!)

Table 6: Free β-chain - EQAS results - arranged according to methods
β-hCG concentration in IU/l (NIBSC 75/551)

Method	N	Sample 1				Sample 2			
		Median	Mean	CV %	SR %	Median	Mean	CV %	SR %
99z	16	15.2	14.8	9.12	94	7.66	7.55	8.87	100

Method	N	Sample 3			
		Median	Mean	CV %	SR %
99z	12	1.20	1.15	8.67	100

Key: As in Table 1 except:
Target values – group median for all samples
Dispersion – ±38% (Current Guideline of the German Medical Council [45])

Calibrated with the 1st international reference preparation (IRP) 1975 for hCG-beta subunit for immunoassay (NIBSC 75/551) 70 IU per ampoule.

Table 7: Calcitonin - EQAS results - arranged according to methods
Calcitonin concentration - ng/l

Method	N	Sample 1				Sample 2			
		Median	Mean	CV %	SR %	Median	Mean	CV %	SR %
2b	21	5770	5900	11.6	90	7.22	7.05	20.3	76
13c	10	4080	4700	17.1	50	9.22	9.47	11.0	80
99z	22	4380	5050	16.9	85	7.30	7.41	16.1	80

Method	N	Sample 3			
		Median	Mean	CV %	SR %
2b	29	<1.00	<1.00		93
13c	14	3.00	2.88	16.2	86
1a/9b/14c/15c	27	<5.00	<5.00		100
99z	7	12.2	12.0	25.2	100

Key: As in Table 1 except:
Target values – group median or less than the values shown (Sample 3 with "<")
Dispersion allowed – target value ±45%

Table 8: Thyroglobulin - EQAS results - arranged according to methods
Thyroglobulin concentration - µg/l

Method	N	Sample 1				Sample 2			
		Median	Mean	CV %	SR %	Median	Mean	CV %	SR %
3b/3c	27	14.6	15.9	19.5	100	25.7	34.4	16.4	100
16c	5	18.0	19.0	18.2	80	42.8	41.0	11.6	100
1a/9b	36	11.1	10.7	12.8	97	29.4	30.0	11.9	100
5b	8	29.4	29.0	9.40	100	61.8	62.7	10.8	100
99z	8	14.8	14.7	23.0	76	29.8	27.5	20.5	76

Method	N	Sample 3			
		Median	Mean	CV %	SR %
3b	11	4.10	4.02	8.92	100
3c	30	5.12	5.03	21.3	94
16c	13	5.51	5.40	12.5	100
1a	40	2.53	2.60	12.3	98
5b	12	9.70	9.95	10.1	100
2b	5	5.40	5.09	18.5	100
99z	12	4.26	4.31	25.7	75

Key: as in Table 1 except:
Target values – group median for all samples
Dispersion – target value ±45%

e. Calcitonin (CT, hCT)

The assays for calcitonin must be able to recognise very high values as such, so that a high-dose hook effect must be eliminated from hCT-assays. The concentrations of calcitonin seen in the serum of unoperated medullary thyroid caricinoma (MCT) patients is in the ng/l to mg/l range. The physician requires accurate pre- and post-operative values in order to assess the success of the operation. Values "greater than" are of little or no use here. The results from three EQA-samples distributed in 2004 are shown in Table 7.

f. Thyroglobulin (Tg)

The kits must measure precisely at the lowest concentration possible (≤2 µg/l) if they are to be used as tumour markers after total thyroidectomy, where recurrence of thyroglobulin in serum or plasma indicates metastatic

growth. The precision of the kits shown in Table 8 is acceptable at all three levels controlled. Kit 5b measures precisely, but consistently higher than the other methods. This could prove a problem for the physician responsible for monitoring thyroid cancer if the laboratory changed methods - for example from manufacturer 3 to manufacturer 5 - and this information failed to reach the physician.

g. Carcinoembryonic Antigen (CEA)

CEA is generally used in assessing colorectal cancers and liver metastases. The generally accepted cut-off values for CEA are between 1.5 and 5 µg/l, depending upon the kit used.
Table 9 shows the results from different manufacturers for CEA Kits in 2 samples dispatched in 2002. Samples were liquid and were commercially prepared from pooled patient sera specially for Instand e.V.
The median values for each method/method-group varied by up to 85%.

Table 9: Carcinoembryonic antigen - EQAS results - arranged according to methods
CEA concentration - µg/l

Method	N	Sample 1				Sample 2			
		Median	Mean	CV %	SR %	Median	Mean	CV %	SR %
5b	138	8.19	8.14	7.4	100	17.2	17.0	7.0	100
4a	202	10.6	10.6	7.3	99	23.4	23.1	6.4	99
10a	13	9.60	9.43	5.2	100	21.4	21.7	4.6	100
1a	32	12.0	12.1	6.3	100	27.7	28.0	6.3	100
11a	23	9.27	9.28	3.8	100	19.6	19.8	4.2	100
6a	23	8.60	8.59	11.9	95	21.4	21.0	9.8	95
9b	14	8.85	8.62	8.8	100	19.3	19.1	9.3	100
9c	4	12.1	11.6	9.5	100	26.2	25.4	10.7	100
7b	41	10.8	10.7	7.7	100	24.9	25.0	7.8	100
12a/17c	16	14.8	14.7	8.2	100	31.7	31.3	12.8	100
3d/15c	7	8.58	8.83	8.3	100	19.0	18.5	8.2	100
99z	2	10.2	10.2	12.5	100	23.6	23.6	19.2	100

Key: As in Table 1 except:
Target values – group median for all samples
Dispersion – target value ±36%

Table 10: CA 15-3 - EQAS results - arranged according to methods
CA 15-3 concentration - kU/l

Method	N	Sample 1				Sample 2			
		Median	Mean	CV %	SR %	Median	Mean	CV %	SR %
4a	159	29.4	29.3	8.1	100	57.0	56.9	8.6	100
5b	83	35.4	35.2	7.7	100	71.0	71.0	9.6	98
9b/9c	22	32.2	32.7	8.8	100	62.4	63.0	7.7	100
7b	28	31.6	32.0	12.9	92	66.7	66.8	8.0	100
11a	13	34.6	34.3	6.1	100	67.9	68.4	7.3	100
15c	4	27.6	28.4	9.4	100	50.1	49.2	12.5	100
3d	5	31.3	30.9	2.3	100	60.8	60.1	2.8	100
1a	21	39.8	39.9	5.3	100	63.8	64.9	10.1	95
6a	11	34.8	34.6	3.2	91	66.5	68.3	7.0	91

Key: As in Table 1 except:
Target values – group median for all samples
Dispersion – target value ±24%

h. CA 15-3

CA 15-3 is used in the in-vitro monitoring of breast cancer. The generally accepted cut-off value for CA 15-3 is 25 kU/l, independent of the kit used.
Table 10 shows the results for CA 15-3 for the same samples as in Table 9.
The inter-kit comparison of the distribution of results was comparable for both samples (variation of the method-median: Sample 1 - 44%; Sample 2 - 42%).

i. CA 19-9

CA 19-9 is often used in conjunction with CEA in monitoring intestinal cancer. The generally accepted cut-off value for CA 19-9 is 37 kU/l.
Table 11 shows the results for CA 19-9 for the same samples as in Table 9. Figure 1 shows a box and whisker plot for Sample 2 in Table 11.
The inter-kit comparison was not comparable for both samples (variation of the method-median: Sample 1 - 116%; Sample 2 - 140%). The method comparison results for CA 19-9 was worse than for CA 15-3.

j. CA 125

CA 125 was found to be useful in monitoring ovarian cancer. The generally accepted cut-off level for CA 125 is 35 kU/l.
Table 12 shows the results for CA 125 for the same samples as in Table 9.
The inter-kit comparison was again not comparable for both samples (variation of the method-median: Sample 1 - 162%; Sample 2 - 108%). The inter-kit variation was similar to that for CA 19-9.

k. CA 72-4

CA 72-4 (formerly known as Tennessee Antigen (TAG)) is used in the follow-up of gastric cancer. The generally accepted cut-off level for CA 72-4 (TAG 72-4) is 4 kU/l.
Table 13 shows the results for CA 72-4 (TAG 72-4) for the same samples as in Table 9.
The inter-kit comparison was comparable for both samples (variation of the method-median: Sample 1 - 445%; Sample 2 - 576%). The variation of results from different kits was the greatest here, although one kit (5b)

Table 11: CA 19-9 - EQAS results - arranged according to methods
CA 19-9 concentration - kU/l

Method	N	Sample 1				Sample 2			
		Median	Mean	CV %	SR %	Median	Mean	CV %	SR %
4a	175	27.8	28.1	11.1	99	60.9	62.3	10.1	99
5b	106	20.9	21.0	10.6	100	43.2	43.2	8.4	96
9b/9c	18	36.0	35.1	13.2	100	83.6	81.7	8.7	100
3d	5	32.0	32.1	4.8	100	59.5	58.8	3.8	100
6a	19	25.1	25.8	14.1	85	64.4	66.4	12.0	85
11a	17	33.8	34.0	11.7	100	84.4	85.5	8.3	100
1a	24	23.6	24.2	12.5	92	57.8	56.4	15.1	92
12a	7	21.2	20.7	5.9	100	35.4	34.8	6.2	100
7b	25	45.9	43.8	12.7	100	104	103	11.9	100

Key: as in Table 1 except:
Target values – group median for all samples
Dispersion – target value ±36%
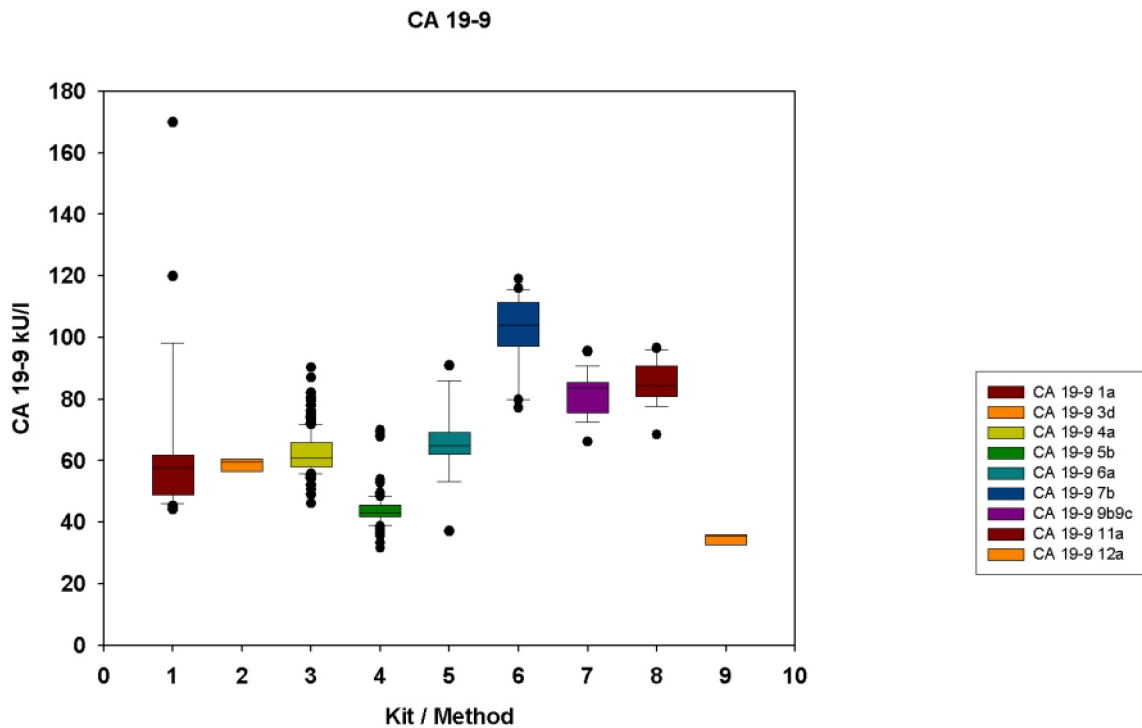
Figure 1: Spread of results for Sample 2 in Tab. 11 listed according to kit/method
The kit/method is shown on the abscissa and in the legend, the CA 19-9 concentrations on the ordinate.
The values between the lower and upper quartiles (25th-75th centiles) are within the box. The whiskers represent the limits ±1.5 x (75th -25th centile values). Outliers are shown as filled circles. The median is shown by the horizontal line within the box.
The number of participants for each kit/method is given in Tab. 11.

gave much lower results that the other kits. Even another kit offered by the same manufacturer (5a) gave results between 6 and 7 times higher than kit 5b. This reflects the absence of standardisation in the determination of the "CA"- tumour markers as a whole, with perhaps the exception of CA 15-3.

I. Alpha-Foetoprotein (AFP)

Alpha foetoprotein has mainly been used for monitoring primary hepatic cancer. The accepted range for healthy non-pregnant individuals is ≤7 kIU/l (≈10 µg/l) (calibrated with NIBSC 72/225).

Table 14 shows the results for AFP for the same samples as in Table 9.
The inter-kit comparison of results is much better than for the "CA"-markers, the variation of the kit-medians for Sample 1 being 25% and for Sample 2, 22%. This may be due to the fact that all AFP kits have been calibrated against the WHO Standard (NIBSC 72/225), which is no longer available.

Table 12: CA 125 - EQAS results - arranged according to methods
CA 125 concentration - kU/l

Method	N	Sample 1				Sample 2			
		Median	Mean	CV %	SR %	Median	Mean	CV %	SR %
4a	142	54.9	55.4	10.5	99	131	131	10.7	99
5a	3	92.8	96.3	8.9	100	203	210	9.7	100
5b	74	50.0	50.2	8.2	100	120	120	8.1	100
9b	24	59.4	58.9	7.8	95	147	145	9.9	100
6a	11	45.4	46.4	6.1	91	123	123	5.8	91
1a	25	38.7	39.0	6.7	96	97.5	98.7	7.2	96
11a	10	39.4	38.8	8.6	100	102	102	8.7	100
9c/15c	6	35.4	35.6	9.0	100	141	141	7.3	100
9b/12a	23	68.9	68.7	9.7	100	172	174	9.3	100

Key: as in Table 1 except:
Target values – group median for all samples
Dispersion – target value ±36%

Table 13: CA 72-4 - EQAS results - arranged according to methods
CA 72-4 concentration - kU/l

Method	N	Sample 1				Sample 2			
		Median	Mean	CV %	SR %	Median	Mean	CV %	SR %
5a	6	18.0	17.2	14.1	100	37.2	36.5	10.2	100
5b	59	3.30	3.25	14.2	98	5.50	5.39	15.3	98
4c/9c	9	11.8	12.2	16.4	100	20.4	21.2	18.5	89
9b	7	12.0	11.1	19.2	100	18.1	18.4	18.2	100
8c/15c/18c	10	14.5	14.0	16.1	100	26.0	25.1	19.9	100
99z	3	3.60	3.48	9.7	100	5.90	6.00	14.2	100

Key: as in Table 1 except:
Target values – group median for all samples
Dispersion – target value ±36%

Table 14: Alpha-foetoprotein - EQAS results - arranged according to methods
AFP concentration - µg/l

Method	N	Sample 1				Sample 2			
		Median	Mean	CV %	SR %	Median	Mean	CV %	SR %
4a	141	22.3	22.4	13.8	88	62.0	61.4	11.3	94
6a	14	21.1	21.1	14.2	93	58.0	57.5	12.7	93
1a	25	22.3	21.8	10.1	96	58.4	58.3	10.7	96
10a	9	21.0	21.4	7.8	100	58.5	58.7	6.4	100
7b	32	26.1	25.6	9.5	97	69.8	69.1	9.0	97
3d/9c/17c/19d/12a	27	22.8	22.4	12.3	96	61.6	60.0	12.6	96
11a	11	20.8	21.8	8.7	100	57.9	59.7	8.5	100
5b	85	22.1	21.8	14.5	96	59.9	59.5	14.0	96
99z	7	24.2	25.8	11.4	72	64.6	66.6	17.9	72

Key: As in Table 1 except:
Target values – group median for all samples
Dispersion – target value ±24% (Current Guideline of the German Medical Council [45])

m. Total Prostate-Specific Antigen (t-PSA); Free PSA (f-PSA); Complexed PSA (c-PSA)

The various forms of circulating prostate-specific antigen (PSA) have been widely used - both singly and in combination - to monitor and differentiate between benign and malignant disorders of the prostate. The generally accepted cut-off for t-PSA is 4 µg/l. The ratios between t-PSA and f-PSA or c-PSA and f-PSA are strictly method-dependent. t-PSA assays may be calibrated with the first international standard (NIBSC 96/670), a mixture of c-PSA (90%) and f-PSA (10%).

Table 15 shows the results for t-PSA, Table 16 for f-PSA and Table 17 for c-PSA in the same samples as for Table 9.

The difference in specificity of the antibody-pairs for PSA determination can be seen in Tables 15 and 16 . The material used for spiking was from seminal fluid, known to be mainly composed of free PSA. Whereas kit 5b differentiated well between "total" and "free" PSA, kit 6a measured more free PSA than total PSA! The results in Table 17 shows the amount of PSA-α_1-antichymotrypsin

Table 15: Total PSA - EQAS results - arranged according to methods
PSA concentration - µg/l

Method	N	Sample 1				Sample 2			
		Median	Mean	CV %	SR %	Median	Mean	CV %	SR %
4a	158	12.6	12.8	10.3	97	3.79	3.80	11.4	97
4b	18	17.7	17.7	4.7	100	5.07	5.05	3.2	100
6a	33	15.5	15.7	9.1	97	4.87	4.85	6.5	97
5b	126	17.3	17.1	8.0	100	5.24	5.23	7.4	100
8a/10a/15c	23	20.7	21.2	11.0	100	6.46	6.65	12.7	100
1a	31	13.9	14.0	9.0	100	4.35	4.31	12.0	100
9b/12a	23	12.5	12.7	11.6	89	4.00	4.00	9.3	93
7b/3d	32	15.9	16.1	8.0	100	4.71	4.74	9.4	100
11a	17	8.80	8.63	8.6	100	3.11	3.09	8.3	100
99z	9	15.0	15.8	18.3	67	4.57	4.59	15.4	67

Key: as in Table 1 except:
Target values – group median for all samples
Dispersion – target value ±31% (Current Guideline of the German Medical Council [45])

Table 16: Free PSA - EQAS results - arranged according to methods
Free PSA concentration - µg/l

Method	N	Sample 1				Sample 2			
		Median	Mean	CV %	SR %	Median	Mean	CV %	SR %
4a/4b	67	11.8	11.9	12.8	97	3.80	3.81	7.7	97
6a	12	19.1	20.6	9.7	92	6.83	6.97	7.1	92
5b	61	8.52	8.49	7.8	100	2.62	2.61	7.2	100
9b	6	8.97	9.49	13.4	84	3.44	3.48	6.6	100
10a	4	15.1	14.9	8.1	100	4.74	4.77	5.5	100
1a	20	9.51	9.33	8.1	100	2.90	2.89	9.7	100
7b/3d/12a	11	13.9	13.4	7.5	100	4.15	4.20	10.8	100
99z	8	11.3	11.3	23.5	75	3.81	3.57	21.3	88

Key: As in Table 1 except:
Target values – group median for all samples
Dispersion – target value ±31% (Current Guideline of the German Medical Council [45])

Table 17: Complexed PSA - EQAS results - arranged according to methods
Complexed PSA concentration as µg/l PSA (not µg/l PSA-ACT complex)

Method	N	Sample 1				Sample 2			
		Median	Mean	CV %	SR %	Median	Mean	CV %	SR %
7b	23	2.26	2.27	6.1	100	0.70	0.69	9.9	100

Key: as in Table 1 except:
Target values – group median for all samples
Dispersion – target value ±31% (Current Guideline of the German Medical Council [45])

(PSA-ACT) in each sample, which is usually much more abundant than free PSA and reflects the non-physiological nature of both samples, due to the reason stated above. Figures 2 and 3 show the results from Tables 15-17 in the form of a box-and whisker plot.

There were large variations in the median values from both total and free PSA kits as can easily be seen in Figures 2 and 3 .

Discussion

The results from the EQA-surveys show that the standardisation of assays used as tumour markers is in many cases far from being optimised. The data presented show that at the present time the hormone assays used for tumour monitoring are on the whole more precise and give rise to more comparable results than the less-specific mucin-carbohydrate "CA-markers" and the various circulating forms of PSA.

The problems of tumour marker measurement have both a physiological and methodological component. For example the determination of CA 19-9, a marker related to the Lewis-antigens [29], [30], [31] is not produced in patients who are both Lewis a and Lewis b negative [31]. This means, that such patients can have large intestinal tumours, which are "negative" for CA 19-9 in serum. Many tumour markers are influenced by renal function, so that dialysis patients and those with impaired renal function may have different (mostly elevated) concentrations when compared with patients with normal renal function [32], [33]. Others, such as AFP, are increased in pregnancy. There are very few "specific" tumour markers. Examples are calcitonin (CT) in medullary thyroid cancer (MTC) and
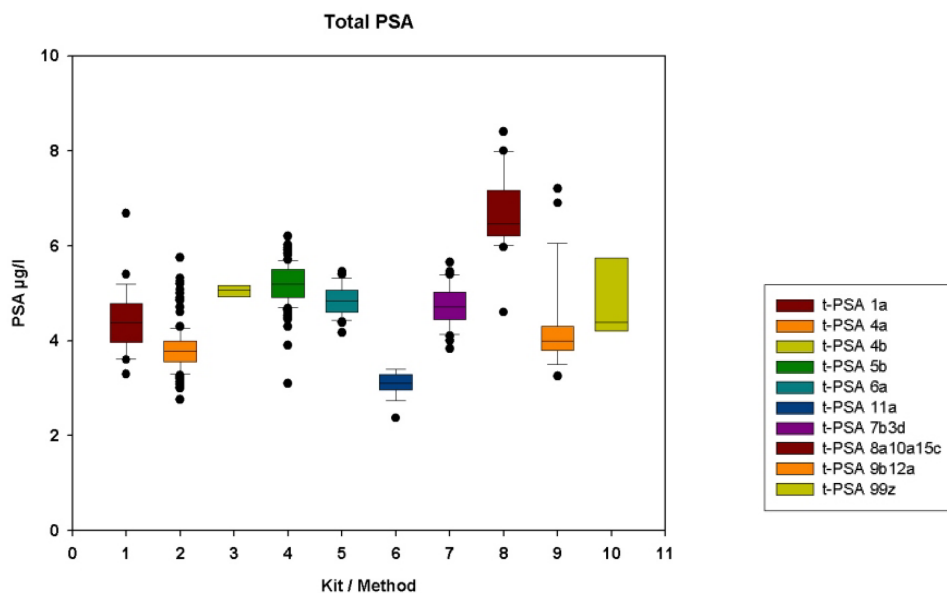
Figure 2: Spread of results for total PSA
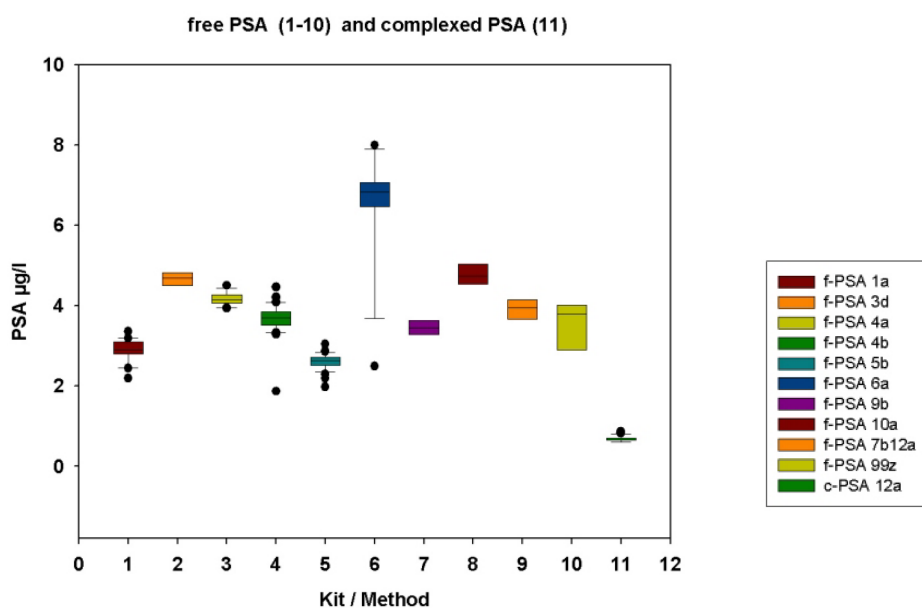The layout for Fig. 2 is identical to that in Fig. 1.

Figure 3: Spread of results for free (10 kits/methods) and complexed PSA (1 kit)
The layout for Fig. 3 is identical to that in Fig. 1.

thyroglobulin (Tg) in thyroidectomised patients, although the latter may be masked by the presence of circulating antibodies to Tg. Elevated levels of chorionic gonadotropin (CG) is relatively specific in males as a marker of germinal-cell cancer.

The commercial interest in tumour-markers cannot be ignored, both from the side of kit-producers as well as from kit-users. The clinical use is far more restricted, many "tumour markers" not being able to fulfil their specifications with regard to both analytical and diagnostic sensitivity and specificity. Clinical decisions made purely on levels of tumour markers - with the exception of perhaps CT, Tg and CG in males - must be seen as irresponsible, especially in a decentralised health system with the free choice of analytical laboratory and methods used.

The EQA results clearly show that the comparability of results and continuity in monitoring patient progress is only possible when using the same method with the same components over a long period of time. The danger of "rationalisation" and "cost-saving" (= buying the cheapest kit offered) practiced by many administrators, coupled with the ignorance of the medical staff on the methodological problems mentioned above, further limits the quality of data obtained from analysing "tumour markers".
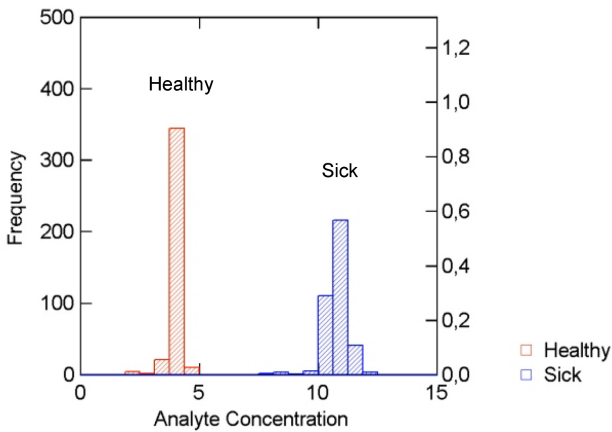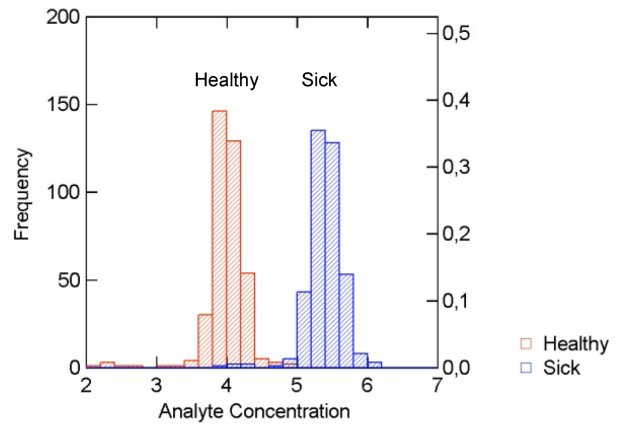
Figure 4: Data for setting diagnostic sensitivity and specificity as histograms
Fig. 4 shows the ideal situation where healthy (red) and sick (blue) patients are clearly separated from each other. The abscissa shows the analyte concentration, the left ordinate the frequency in each column. The right ordinate shows the relative frequency distribution (Value: 0=0%; 1=100%).
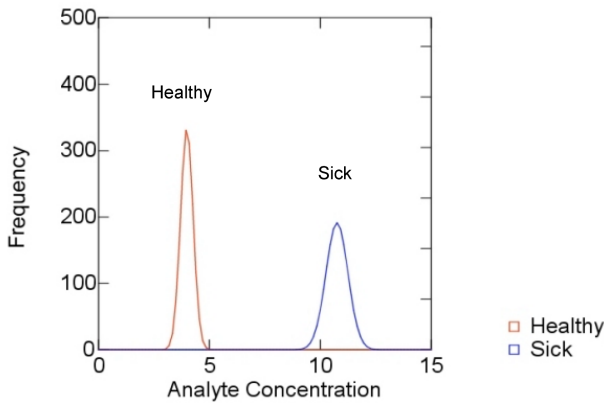
Figure 5: Smoothed data distribution curves for the data in Fig. 4
The abscissa and left ordinate are as in Fig. 4.

The effect of numerical values can be visualised in comparing the diagnostic sensitivity (the correct prediction of a positive (=tumour present) result) and diagnostic specificity (the correct prediction of a negative (=tumour not present) result). Figures 4 and 5 show the ideal case for a tumour marker, where healthy (red) and sick (blue) patients are clearly separated from each other (diagnostic sensitivity and specificity 100%). Figures 6 and 7 show the more common situation, i.e. where both groups overlap. In Figure 7, there are three ways of setting a decision point. At point A we have the lowest analyte concentration where all sick patients are correctly allocated (diagnostic sensitivity 100%). A number of healthy patients have however concentrations higher than this point (diagnostic specificity <100%) and would be classified as sick. At point B we have the converse situation -

the highest concentration where all healthy individuals are correctly classified (diagnostic specificity 100%), but where some sick individuals (those under the blue curve to the left of B) are to be found (diagnostic sensitivity <100%) and are classified wrongly as being healthy. Point C represents the best compromise where the degree of false classification is minimised, but where sick and healthy individuals in the border region may be wrongly classified (diagnostic sensitivity and specificity in this case <100%).

Figure 6: Data for setting diagnostic sensitivity and specificity as histograms
Fig. 6 shows the real situation where the groups of healthy (red) and sick (blue) patients overlap. The abscissa shows the analyte concentration, the left ordinate the frequency in each column. The right ordinate shows the relative frequency distribution (Value: 0=0%; 1=100%).
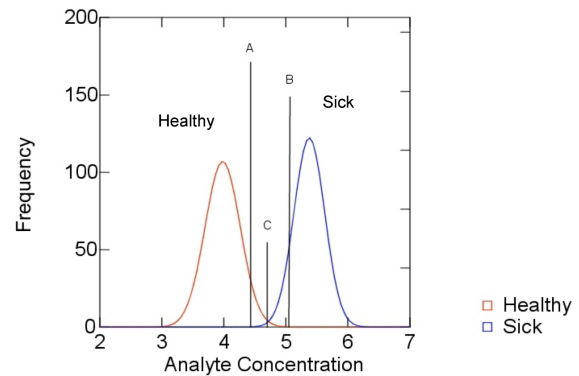

Figure 7: Smoothed data distribution curves for the data in Fig. 6
Fig. 7 shows the analyte concentration for 100% diagnostic sensitivity (A) and 100% diagnostic specificity (C). Point B shows the compromised "real-life" situation, where both diagnostic sensitivity and specificity are less than 100%. The abscissa and ordinates are as in Fig. 6.

In the case of t-PSA, if we leave the points A, B and C in Figure 7 where they are and measure patient sera with the methods in Table 15, it becomes obvious that the red and blue curves will be moved - either to the right or to
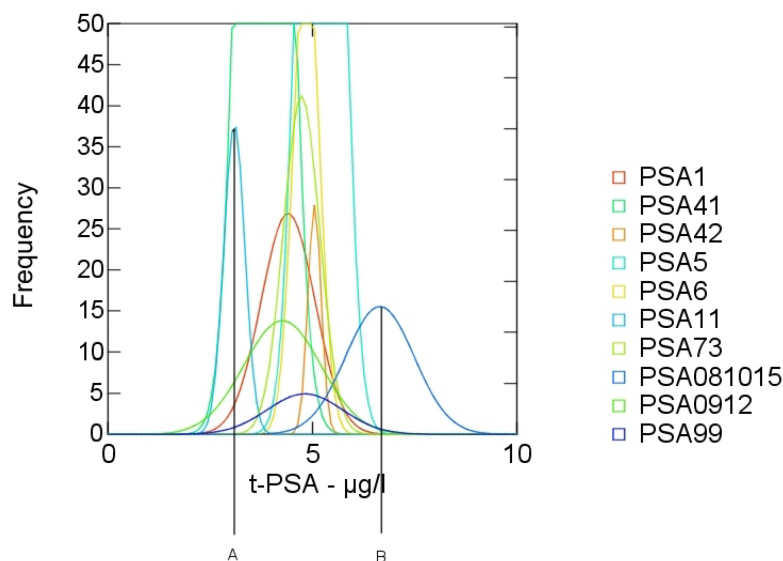
Figure 8: Data for Sample 2 in Tab. 15 (total PSA) highlighting the variabilty of measurement, both in absolute concentration and precision for the different kits on the market
The abscissa and ordinate are as in Fig. 5 and 7.

the left - thus changing the diagnostic sensitivity and specificity according to the method/kit used. The points A,B and C must be evaluated for each kit.

Figure 8 shows the distribution of data of the t-PSA Sample 2 in Table 15. Both the mean values and precision of measurement (seen by the different degrees of kurtosis) make it clear that the variability in measurement - here taking t-PSA as the example - must lead to different kit-specific cut-off values, thus nullifying a static point of decision at 4 µg/l. Both extremes are shown by the points A (Kit 11a) and B (Kits 8a, 10a and 15c).

Standardisation of methods - not with international antigen preparations but with defined antigenic epitopes recognised by defined monoclonal antibodies, as in the case of the tumour marker Cyfra 21-1 [34], can lead to more comparability between manufacturers and to an improvement in the interpretation of results by clinical staff, thus improving the benefits to patients subjected to such analyses.

Method-dependent differences in results, especially in the case of PSA, make it imperative that method-specific reference ranges and "cut-off" values must be established for each analyte. In the case of PSA, where ratios between free and total PSA or complexed and total PSA are used in differentiating between benign and malignant disease, kits from the same manufacturer must be used for free and total/complexed PSA, if the ratios given by the manufacturer are to have any clinical use. General reference-ranges and ratios often given in text books are of no use, due to the individual nature of results from different manufacturers, although some authors [35] point out the importance of using kit-specific reference ranges. The importance of this is shown in Tables 15 and 16 above, even though the samples were not physiological (concentration of free PSA much higher than complexed PSA). Even kits from the same manufacturer, but developed for different instruments with different measuring tech-

niques may give rise to statistically different results from the same sample - here seen in kits 41 and 42 for t-PSA. The use of generalised figures for the diagnostic sensitivity and specificity of tumour-marker kits must be discouraged, as seen in the results from this study (as an example, see Figure 1). Each laboratory must therefore estimate its own diagnostic sensitivity and specificity for all tumour-markers from data obtained from the kit used and the group(s) of patients studied. The relatively low diagnostic sensitivity and/or specificity shown by many tumour-markers reflects their inability to be used for screening purposes. Despite this, medical staff still use tumour markers for patient screening - especially PSA - although no prospective randomised trial has been performed to validate or reject the use of PSA in screening for prostate cancer [36]. Recently, the value of PSA in diagnosis and monitoring of prostate cancer has again been questioned [37]. Even in review articles, the cut-off for t-PSA decision taking is set at 4 µg/l, despite evidence for method dependent reference ranges cited in the same article [38]!

Even in the case of post-operative follow up, it is possible, that when metastasis occurs, "tumour-specific-antigens" may be produced which are structurally different from those in the primary tumour, thus rendering the further use of the original marker for follow-up useless.

The future of detection and post-operative follow-up of tumours may very well lie in the developing field of gene-expression chips [39]. That synthesis of tumour markers can change during the development of neoplasms has been shown in immunohistochemical studies [40].

Even the use of genetical markers can at present only predict a predisposition for - in this case - the development of a certain tumour. The time of appearance of the lesion cannot be predicted from the presence of or mutations in certain genes - for example BRCA-1 [24],

which usually lead to the development of breast and/or ovarian cancer before the 5th decade of life.

Classical methods of detection and control, such as x-ray, sonography, PET, MRT and computer tomography, will remain the "golden" standards of tumour detection and control of tumour progression/remission for many years to come, although "gene-chips" will play an increasingly important role in detection of those at risk, as well as post-operative follow-up of cancer patients [41], [42], [43], [44].

To conclude, the majority of analytes used as tumour markers are unsuitable for screening purposes. Exceptions are the measurement of calcitonin, chorionic gonadotropin and thyroglobulin when used in the situations stated above. The use of the other analytes reported in this article may be of use in longitudinal follow-up pre- and post-operatively, with the restriction that the method used for a given analyte must always be the same, at least until both an acceptable standardisation of methods/results and comparability of numerical values (concentrations) is achieved. The latter has not been realised and is not to be expected in the near future, at least for the majority of analytes. Until then, defined individual cut-off values must be established for each tumour marker kit on the market.

References

1. von Euler U, Lishajko F. Improved technique for the fluorimetric estimation of catecholamines. Acta Physiol Scand. 1961;51:348-55.

2. Turler K, Kaser H. Quantitative fluorimetric determination of urinary dopa and its significance for the diagnosis of neural crest tumours. Clin Chim Acta. 1971;32(1):41-51.

3. Wood WG, Mainwaring-Burton RW. The development and evaluation of a semi-automated assay for catecholamines suitable for plasma and urine. Clin Chim Acta. 1975;61(3):297-308.

4. Feldman JM. Serotonin metabolism in patients with carcinoid tumors: incidence of 5-hydroxytryptophan-secreting tumors. Gastroenterology. 1978;75(6):1109-14.

5. Huggins C, Hodges CV. Studies on prostatic cancer. The effects of castration, of estrogen and of androgen injection on serum phosphatases in metastatic carcinoma of the prostate. Cancer Res. 1941;1:293-7.

6. Berson SA, Yalow RS. Radioimmunoassay of ACTH in plasma. J Clin Invest 1968;47(12):2725-51.

7. Bloomfield GA, Holdaway IM, Corrin B, Ratcliffe JG, Rees GM, Ellison M, Rees LH. Lung tumours and ACTH production. Clin Endocrinol (Oxf). 1977;6(2):95-104.

8. Rushworth AG, Orr AH, Bagshawe KD. The concentration of HCG in the plasma and spinal fluid of patients with trophoblastic tumours in the central nervous system. Br J Cancer. 1968;22(2):253-7.

9. Tashjian AH Jr, Melvin EW. Medullary carcinoma of the thyroid gland. Studies of thyrocalcitonin in plasma and tumor extracts. N Eng J Med. 1968;279(6):279-83.

10. Murphy BE. Radioassays of non-antigenic hormones. Trans Assoc Am Physicians. 1968;81:92-103.

11. Midgley AR Jr, Niswender GD, Ram JS. Hapten-radioimmunoassay: a general procedure for the estimation of steroidal and other haptenic substances. Steroids 1969;13(6):731-7.

12. Charters AC, Odell WD, Thompson JC. Anterior pituitary function during surgical stress and convalescence. Radioimmunoassay measurement of blood TSH, LH, FSH and growth hormone. J Clin Endocrinol Metab. 1969;29(1):63-71.

13. Gold P, Freedman SO. Demonstration of tumor specific antigens in human colonic carcinoma by immunologic tolerance and absorption technique. J Exp Med. 1965;121:439-62.

14. Krupey J, Gold P, Freedman SO:Physicochemical studies of the carcinoembryonic antigens of the human digestive system. J Exp Med. 1968;128(3):387-98.

15. Del Villano BC, Zurawski VR Jr. The carbohydrate antigenic determinant 19-9 (CA 19-9):a monoclonal antibody defined tumor marker. Lab Res Methods Biol Med. 1983;8:269-82.

16. Gang Y, Adachi I, Ohkura H, Yamamoto H, Mizuguchi Y, Abe K. CA 15-3 is present as a novel tumour marker in the sera of patients with breast cancer and other malignancies. Gan To Kagaku Ryoho. 1985;12(12):2379-86.

17. Bast RC Jr, Feeney M, Lazarus H, Nadler LM, Colvin RB, Knapp RC. Reactivity of a monoclonal antibody with human ovarian carcinoma. J Clin Invest. 1981;68(5):1331-7.

18. Ohuchi N, Takahashi K, Matoba N, Mori S. A radioimmunometric assay for circulating tumor-associated glycoprotein (TAG-72) using the antigen determinant CA 72-4. Gan To Kagaku Ryoho. 1988;15(9):2761-6.

19. Wood WG, van der Sloot E, Böhle A. The establishment and evaluation of luminescent-labelled immunometric assays for prostate-specific antigen-alpha 1-antichymotrypsin complexes in serum. Eur J Clin Chem Clin Biochem. 1991;29(12):787-94.

20. Wu JT, Wilson LW. Development of a microplate ELISA for free PSA and PSA-ACT complex in serum. J Clin Lab Anal 1995;9(4):252-60.

21. Mitrunen K, Pettersson K, Piironen T, Bjork T, Lilja H, Lövgren T. Dual-label one-step immunoassay for simultaneous measurement of free and total prostate-specific antigen concentrations and ratios in serum. Clin Chem. 1995;41(8 Pt 1):1115-20.

22. Stenman UH, Leinonen J, Alfthan H, Rannikko S, Tuhkanen K, Alfthan O. A complex between prostate-specific antigen and alpha1-antichymotrypsin is the major form of prostate-specific antigen in serum of patients with prostatic cancer:assay of the complex improves clinical sensitivity for cancer. Cancer Res. 1991;51(1):222-6.

23. Vessella RL, Lange PH. Issues in the assessment of PSA immunoassays. Urol Clin North Am. 1993;20(4):607-19.

24. Hall JM, Freedman L, Guenther C, Lee MK, Weber JL, Black DM, King MC. Closing in on a breast cancer gene on chromosome 17q. Amer J Hum Genet. 1992;50(6):1235-42.

25. Schutte M, Rozenblum E, Moskaluk CA, Guan X, Hoque AT, Hahn SA, da Costa LT, de Jong PJ, Kern SE. An integrated high-resolution physical map of the DPC/BRCA2 region at chromosome 13q12. Cancer Res. 1995;55(29):4570-4.

26. Cornelisse CJ, Cornelis RS, Devilee P. Genes responsible for familial breast cancer. Pathol Res Pract. 1996;192(7):684-93.

27. Stieber P, Thomas L, Lamerz R, Raue F, Grauer A, Nollau P et al. Tumormarker (Kapitel 28). In: Thomas L, ed. Labor und Diagnose - Indikation und Bewertung von Laborbefunden für die medizinische Diagnostik. 6. Aufl. Frankfurt/Main: TH-Books Verlagsgesellschaft mbH; 2005. p. 1291-1366.

28. Thomas L, Lüthgens M, Wagener C, Lamerz R, Mann K, Tatra G et al. Tumormarker. In: Thomas L, ed. Labor und Diagnose - Indikation und Bewertung von Laborbefunden in die medizinische Diagnostik. 3., überarb. u. erw. Aufl. Marburg: Med. Verlagsgesellschaft; 1988. p. 945-1015.

29. Herrero-Zabaleta ME, Gautier R, Burtin P, Daher N, Bara J. Monoclonal antibody against sialylated Lewis(a) antigen. Bull Cancer. 1987;74(4):387-96.

30. Hirohashi S, Shimosato Y, Ino Y, Tome Y, Watanabe M, Hirota T, Itabashi M. Distribution of blood group antigens and CA 19-9 in gastric cancers and non-neoplastic gastric mucosa. Gann. 1984;75(6):540-7.

31. Pour PM, Tempero MM, Takasaki H, Uchida E, Takiyama Y, Burnett DA, Steplewski Z. Expression of blood group related antigens ABH, Lewis A, Lewis B, Lewis X, Lewis Y, and CA 19-9 in pancreatic cancer cells in comparison with the patient's blood group type. Cancer Res. 1988;48(19):5422-6.

32. Wood WG, Steinhoff J, Kessler AC. Anomalous tumour marker concentrations in renal transplant patients. Eur J Clin Chem Clin Biochem. 1993;31(2):75-82.

33. Wood WG, Steinhoff J, Gutekunst R. Sense and nonsense in determination of tumour markers in serum - results with potentially disastrous consequences. In: Klapdor R, ed. Tumor associated antigens, oncogenes, receptors, cytokines in tumor diagnosis and therapy at the beginning of the nineties. München: Zuckschwerdt; 1992. p. 305-12.

34. Bodenmüller H, Ofenloch-Hähnle B, Lane EB, Dessauer A, Böttger V, Donie F. Lung cancer-associated keratin 19 fragments:development and biochemical characterisation of the new serum assay Enzymun-Test CYFRA 21-1. Int J Biol Markers. 1994;9(2):75-81

35. Lamerz R. PSA (Prostata-spezifisches Anitigen). In: Thomas L, ed. Labor und Diagnose. 5. ed. Frankfurt/Main: TH-Books Verlagsgesellschaft; 1998. p. 1004-8 (Kap. 34.14).

36. Duffy MJ. Evidence for the clinical use of tumour markers. Ann Clin Biochem. 2004;41(Pt 5):370-7.

37. Stamey TA, Caldwell M, McNeal JE, Nolley R, Hemenez M, Downs J. The prostate specific antigen era in the United States is over for prostate cancer: what happened in the last 20 years? J Urol. 2004;172(4 Pt 1):1297-301.

38. Brawer MK. Prostate-specific antigen: current status. CA Cancer J Clin. 1999;49(5):264-81.

39. Mutter GL, Baak JP, Fitzgerald JT, Gray R, Neuberg D, Kust GA, Gentleman R, Gullans SR, Wei LJ, Wilcox M. Global expression changes of constitutive and hormonally regulated genes during endometrial neoplastic transformation. Gynecol Oncol. 2001;83(2):177-85.

40. Van Niekerk CC, Boerman OC, Ramaekers FC, Poels LG. Marker profile of different phases in the transition of normal human ovarian epithelium to ovarian carcinomas. Am J Pathol. 1991;138(2):455-63.

41. Holleman A, Cheok MH, den Boer ML, Yang W, Veerman AJ, Kazemier KM, Pei D, Cheng C, Pui CH, Relling MV, Janka-Schaub GE, Pieters R, Evans WE. Gene-expression patterns in drug-resistant acute lymphoblastic leukemia cells and response to treatment. N Eng J Med. 2004;351(6):533-42.

42. Morrison C, Farrar W, Kneile J, Williams N, Liu-Stratton Y, Bakaletz A, Aldred MA, Eng C. Molecular classification of parathyroid neoplasia by gene expression profiling. Am J Pathol. 2004;165(2):565-76.

43. Inazawa J, Inoue J, Imoto I. Comparative genomic hybridisation (CGH)-arrays pave the way for identification of novel cancer-related genes. Cancer Sci. 2004;95(7):559-63.

44. Whipple ME, Mendez E, Farwell DG, Agoff SN, Chen C. A genomic predictor of oral squamous cell carcinoma. Laryngoscope. 2004;114(8):1346-54.

45. Bundesärztekammer. Richtlinie der Bundesärztekammer zur Qualitätssicherung quantitativer laboratoriumsmedizinischer Untersuchungen. Gemäß Beschluss des Vorstandes der Bundesärztekammer vom 24. August 2001, veröffentlicht im Deutschen Ärzteblatt 98, Heft 42, 19. Oktober 2001, Seite A 2747 - 2759. Zuletzt geändert durch Beschluss des Vorstandes der Bundesärztekammer vom 14. November 2003, veröffentlicht im Deutschen Ärzteblatt 100, Heft 50, 12. Dezember 2003, Seite A 3335 - 3338. Dtsch Ärztebl. 2003;100(50):A3335-A3338. Available from: http://www.bundesaerztekammer.de/30/Richtlinien/Richtidx/Labor2002/RiliLabor.pdf

Corresponding author:

Prof. William Graham Wood
Instand e.V., Ubierstrasse 20, 40223 Düsseldorf, Germany
wood@instand-ev.de