

**Aus dem Institut für Laboratoriumsmedizin
Klinikum-Bremen-Mitte
(ehemaliger Direktor: Prof. Dr. med. Rainer Haeckel)**

**An diagnostischen Entscheidungen orientierter
Methodenvergleich von POCT-Glukometern für die
Diagnose eines Typ 2 Diabetes mellitus**

**Dissertation
Zur Erlangung des Doktorgrades der Medizin
in der Medizinischen Hochschule Hannover**

**vorgelegt von
Isabel Rosa Christiane Püntmann
aus Fulda**

Hannover 2009

Angenommen vom Senat der Medizinischen Hochschule Hannover
am
15.06.2009

Gedruckt mit Genehmigung der
Medizinischen Hochschule Hannover

Präsident:	Prof. Dr. med. Dieter Bitter-Suermann
Betreuer der Arbeit:	Prof. Dr. med. Rainer Haeckel
Referent:	Prof. Dr. rer. Nat. Eberhard Gurr
Korreferent:	Prof. Dr. med. Gerhard Schumann
Tag der mündlichen Prüfung:	15.06.2009
Promotionsausschussmitglieder:	Prof. Dr. Hans-Heinrich Kreipe Prof. Dr. Sebastian Suerbaum Prof. Dr. Reinhard Brunkhorst

Meinen Eltern in Dankbarkeit gewidmet

INHALTSVERZEICHNIS

1	EINLEITUNG	3
2	FRAGESTELLUNGEN	5
3	MATERIAL UND METHODEN	6
3.1	MATERIAL	6
3.1.1	Auswahl der Glukometer	6
3.1.1.1	Referenzgerät	6
3.1.1.2	Testgeräte	8
3.1.2	Materialien und Reagenzien	12
3.2	METHODEN	13
3.2.1	Standard- und Kontroll-Lösungen	13
3.2.1.1	Mess-Serien, Sollwert-Abweichung, Tages-bezogene Korrektur	15
3.2.2	Biologisches Probenmaterial	16
3.2.2.1	Oraler Glukosetoleranztest	16
3.2.2.2	Gewinnung des Probenmaterials und Analyse der Glukosekonzentration	17
3.2.2.3	Mess-Serien und Vergleiche der verschiedenen Probenarten	19
3.2.2.4	Probanden, Kollektive und Zeiträume	20
3.3	STATISTISCHE AUSWERTUNG	23
3.3.1	Ermittlung der Rate falsch positiver und falsch negativer Messwerte	23
3.3.2	Konventioneller Methodenvergleich	24
3.3.3	Diskordanzraten-Test	27
3.3.3.1	Erforderliche Daten und Berechnungsschritte des Diskordanzraten-Tests	28
3.3.3.2	Graphische schematische Darstellung des Diskordanzraten-Tests	33
3.3.3.3	Anwendung des Diskordanzraten-Tests in der vorliegenden Studie	37
4	ERGEBNISSE	38
4.1	MESSUNGEN DER STANDARD- UND KONTROLL-LÖSUNGEN	38
4.1.1	Sollwert-Abweichungen und Impräzision	38
4.1.2	Tages-bezogene Korrekturfaktoren	41
4.2	MESSUNGEN DES MENSCHLICHEN PROBENMATERIALS	42
4.2.1	Rate an falsch positiven und falsch negativen Messwerten	42
4.2.2	Konventioneller Methodenvergleich	44
4.2.3	Diskordanzraten-Test	48
4.2.3.1	Diskordanzraten der intrasystemischen Vergleiche	48
4.2.3.2	Diskordanzraten der intersystemischen Vergleiche	51

5	DISKUSSION	58
6	ZUSAMMENFASSUNG	66
7	LITERATURVERZEICHNIS	68
8	ANHANG	73
8.1	TABELLEN	73
8.2	ABBILDUNGEN	77
8.3	TABELLENVERZEICHNIS	106
8.4	ABBILDUNGSVERZEICHNIS	108
8.5	ABKÜRZUNGEN	109
8.6	LEBENS LAUF	111
8.7	ERKLÄRUNG	114
9	DANKSAGUNG	116

1 EINLEITUNG

In den letzten Jahren haben Geräte zur patientennahen Diagnostik, dem sogenannten Point-of-Care-Testing (POCT), mehr und mehr Eingang in den klinischen Alltag gefunden. Sie wurden entwickelt, um schnelle und laborunabhängige Diagnostik zu ermöglichen, was in Akutsituationen eine effizientere Therapie und bei chronischen Erkrankungen eine kontinuierliche Therapieeinstellung durch regelmäßiges Monitoring gewährleisten kann. Im Rahmen der Überprüfung neuer Messmethoden mit konventionellen Labormessmethoden gewann infolgedessen der Vergleich von Laboranalysen zunehmend an Bedeutung. Die bei solchen Methodenvergleichen eingesetzten statistischen Testverfahren beurteilen allerdings ausschließlich die analytische Leistungsfähigkeit; sie geben keine Antwort auf die klinisch relevante Frage, wieviel diagnostische Fehlentscheidungen mit einem neuen Verfahren im Vergleich zu einem Routineverfahren getroffen werden.

Beim Vergleich zweier Analyseverfahren werden in der klinischen Chemie klassische Zuverlässigkeitskriterien wie Präzision, Richtigkeit, Interferenzen und Nachweisgrenzen zur Beurteilung herangezogen. Vergleichende Ergebnisse mit Patientenproben werden mittels Regressionsverfahren und Differenzenplot (nach Altman und Bland [1;2]) ausgewertet. Für Glukometer, insbesondere im Einsatz für POCT, wurden unterschiedliche zulässige Fehlerraten von mehreren Autoren und Consensusgruppen angegeben (Tabelle 1 im Anhang), deren zugrundeliegende Überlegungen meistens ungenügend belegt sind. Zwischen den verschiedenen Vorschlägen existiert kein Konsens.

Bei den analytischen Zuverlässigkeitskriterien handelt es sich ausschließlich um deskriptive Größen, die zwar dem Labormediziner eine Gütebeurteilung erlauben, dem Kliniker aber keine Information darüber liefern, mit welchen diagnostischen Konsequenzen er rechnen muss, wenn von einem Verfahren auf ein anderes gewechselt wird. Fraser und Petersen postulieren in ihrer „Objektivitäts-Hierarchie“ („hierarchy of objectivity“) [3] (Tabelle 2 im Anhang), dass vorrangig klinische Erfordernisse in der Beurteilung der analytischen Zuverlässigkeit von Messgeräten berücksichtigt werden sollten. Zuverlässigkeitskriterien und Qualitätsspezifikationen sollten im Hinblick auf den klinischen Entscheidungsprozess und auf den medizinischen Nutzen objektivierbar sein. Experten- und Konsensus-Empfehlungen über die zulässigen Fehlerraten (Tabelle 1 im Anhang) werden in dieser Hierarchie an zweiter Stelle eingestuft.

In den letzten Jahren wurden Konzepte entwickelt, um die klinische Relevanz und die klinische Konsequenz bei Methodenvergleichen beurteilen zu können, um also der höchsten Hierarchiestufe der Objektivität von Fraser und Petersen [3] zu entsprechen. Die von Cox *et*

al. [4] entwickelte error grid Methode und das modifizierte Verfahren von Clarke *et al.* [5] und Koschinsky *et al.* [6], bei dem der gesamte Messbereich in verschiedene Sektoren aufgeteilt wird, versuchen klinische Erfordernisse zu berücksichtigen, führen aber schließlich nur zu einer Prozentangabe für den zulässigen Fehler. Boyd und Bruns [7] haben die Leistungsmerkmale von Glukosemessgeräten in Bezug zu der Fehlerrate bei der Insulindosierung gesetzt. Bei diesem Ansatz ging es um die klinische Relevanz bezüglich der Therapie, nicht aber bezüglich der Diagnose.

Der klassische Ansatz zur Beurteilung der diagnostischen Zuverlässigkeit ist die Ermittlung falsch bzw. richtig positiver und negativer Werte. Die Verfahren zur Bestimmung der diagnostischen Zuverlässigkeit erfordern eine relativ große Datenmenge, die bei den üblichen Methodenvergleichen ($n = 50 - 100$) nicht gegeben ist. In der vorliegenden Studie wird ein neues Verfahren angewandt [8-14], das die zu erwartenden Fehlerraten (Diskordanzen) an der Entscheidungsgrenze schätzt, wenn von einem Analyseverfahren auf ein anderes übergegangen wird. Es lässt sich bereits mit den beim Methodenvergleich üblichen Datenmengen anwenden. Dieses Verfahren liefert die Information, die der Diagnostiker vom Methodenvergleich erwartet. Es beantwortet die Frage, wie viele diskordante Klassifikationen auftreten, wenn zwei Methoden genutzt werden, um einen pathologischen von einem nicht-pathologischen Zustand zu diskriminieren. Basierend auf klinischen Entscheidungen kann dieser entscheidungsorientierte Methodenvergleich, genannt Diskordanzraten-Test, in der Objektivitätshierarchie von Fraser und Petersen [3] der Stufe 1 zugeordnet werden.

Im Rahmen der vorliegenden Studie wurde die analytische Zuverlässigkeit von POCT-Glukometern gegenüber einem etablierten Laboranalysegerät überprüft. Es wurden einerseits Glukosekonzentrationsbestimmungen unter Standardbedingungen mit vordefinierten Standard- und Kontroll-Lösungen durchgeführt. Andererseits wurde die Glukosekonzentration unter klinischen Bedingungen, also in biologischer Matrix (Blut, Plasma) bestimmt, die im Rahmen von oralen Glukosetoleranztests gewonnen wurde. Darüberhinaus wurde anhand der Messergebnisse zweier etablierter Routinegeräte überprüft, ob sich der Diskordanzraten-Test zum Methodenvergleich eignet, und ob die von POCT-Glukometern erzielten Diskordanzraten denen etablierter Routinemessgeräten entsprechen. Daran schloss sich die Frage, welche Konsequenz man aus dem entscheidungsorientierten Methodenvergleich ziehen kann, insbesondere ob POCT-Geräte für die Diagnosestellung eines Diabetes mellitus geeignet sind.

Die vorliegende Studie ist keine Evaluation der analytischen Zuverlässigkeitskriterien, sondern in erster Linie eine Überprüfung der diagnostischen Vergleichbarkeit der ausgewählten Glukometer.

2 FRAGESTELLUNGEN

1. Ist die analytische Zuverlässigkeit der ausgewählten POCT-Glukometer mit etablierten Routineanalysegeräten vergleichbar?
2. Lassen sich POCT-Glukometer mittels wässriger oder proteinhaltiger Kontrollmaterialien auf das laborinterne Routineanalysegerät beziehen?
3. Eignet sich die Methode der Diskordanzraten für den Vergleich etablierter Routineanalysegeräte zur Bestimmung der Glukosekonzentration?
4. Werden die von Routineanalysegeräten erzielten Diskordanzraten auch von POCT-Glukometern erreicht?
5. Welche Konsequenzen können für die Praxis der Diagnostik des Typ 2 Diabetes mellitus abgeleitet werden?

3 MATERIAL UND METHODEN

3.1 MATERIAL

3.1.1 Auswahl der Glukometer

3.1.1.1 Referenzgerät

Als Referenzgerät wurde ein Ebio Plus 6668 Analysesystem [15] (Eppendorf AG Hamburg) gewählt (im fortlaufenden Text als Ebio abgekürzt), welches ein seit Jahren etabliertes Routinelaboranalysegerät des Zentrallaboratoriums des Klinikums Bremen-Mitte (ehemals Zentralkrankenhaus Sankt Jürgenstraße Bremen) war. Das Labor war akkreditiert nach EN 45001 [16;17] und erfüllte alle Qualitätsanforderungen, die in den Richtlinien der Bundesärztekammer [18;19] vorgegeben waren. Die Leistungsfähigkeit wurde entsprechend der Anforderungen der ECCLS-Leitlinien (European Committee of Clinical Standards) [20] im Rahmen multizentrischer Evaluationen nachgewiesen [21]. Qualitätskontrollen erfolgten extern im Rahmen von bundesweit durchgeführten Ringversuchen und intern durch tägliche Standard- und Kontroll-Messungen. Ebio war demnach ein sorgfältig evaluiertes und qualitätsgesichertes Routinelaboranalysesystem, dessen Messgenauigkeit der vorgesehenen Verwendung entsprach. Für die Anwendung als Referenzgerät in Methodenvergleichen gegenüber anderen Glukometern bei der Glukosebestimmung in Vollblut und Plasma war es im höchsten Maße geeignet.

Die Glukosekonzentration wurde in einer 10 µl-Probe (Standardlösung, Blut oder Plasma) in heparinisierten End-zu-End Glaskapillaren + 500 µl hämolysierendem Reagenz mittels einer Glukoseoxidase-enthaltenden ionensensitiven Elektrode nach enzymatisch-amperometrischem Messprinzip bestimmt [15;21].

Ein Messzyklus beginnt mit dem Eintauchen der Probenkanüle in eine Probe, wobei Probenlösung zusammen mit dem Hämolysereagenz in die Messkammer gesaugt wird (Messprinzip des Ebio, Abbildung 1). Die bei der Probennahme in die Messkammer gelangte Glukose trifft nach Durchtritt durch die erste Membrangrenzschicht auf die immobilisierte Glukoseoxidase und wird mit Hilfe der katalytischen Wirkung zu Glukonsäure und Wasserstoffperoxid (H_2O_2) umgesetzt [22]. Nach der Diffusion durch die zweite Membrangrenzschicht wird H_2O_2 an der Platinelektrode bei +600 mV oxidiert:

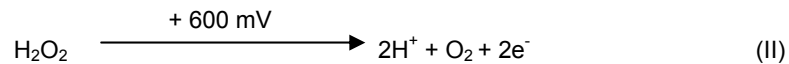
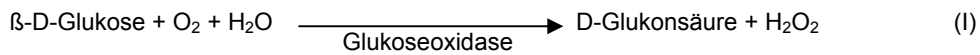
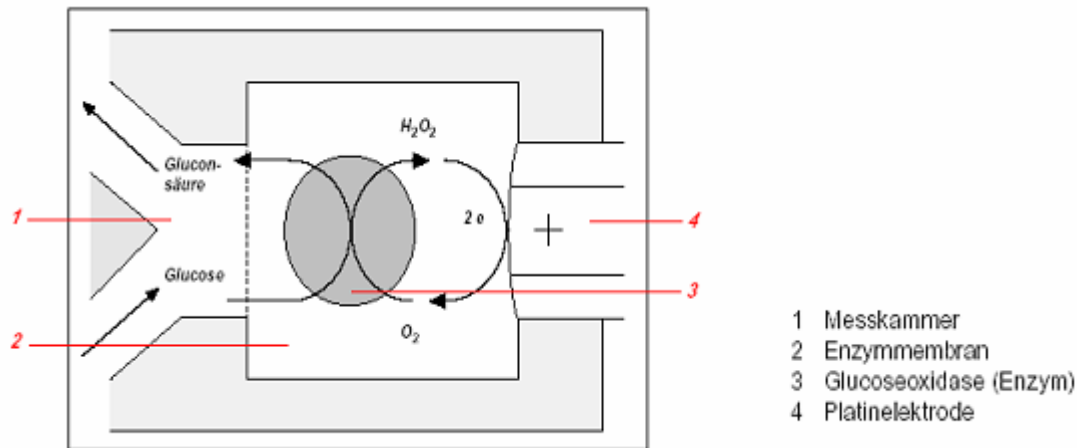
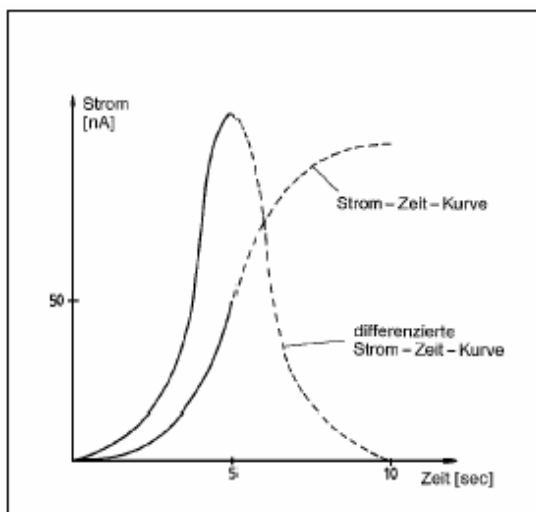


Abbildung 1: Messprinzip des Ebio [15]



Das so erhaltene Messsignal (Strom-Zeit-Kurve, Abbildung 2) wird während der Probennahme fortlaufend differenziert. Das Maximum der differenzierten Kurve kennzeichnet den maximalen Anstieg der Strom-Zeit-Kurve. Das hierzu gehörige Messsignal (Maximum der differenzierten Kurve) wird in einen Spannungswert umgewandelt. Dieser Messwert ist der Glukosekonzentration der Probe proportional.

Abbildung 2: Strom-Zeit-Kurve [15]



Das Ebio-Analysegerät kann sowohl Vollblut-, Plasma-, Serum- als auch Kapillarblut-Proben analysieren. Alle Proben werden grundsätzlich mit der gleichen Hämolyseflüssigkeit verdünnt. Der Messbereich für Glukose liegt nach Angaben des Herstellers zwischen 11 und 900 mg/dl.

3.1.1.2 Testgeräte

Die Auswahl der Vergleichs- bzw. Testgeräte richtete sich danach, welche POCT-Geräte während des Studienzeitraums im Klinikum Bremen-Mitte in Verwendung waren. Als portable POCT-Glukometer wurden die handflächengroßen Glukometer Accutrend Sensor [23;24], Accu-Chek Plus (zwei Geräte) von Roche Diagnostics GmbH Mannheim [25], Glucometer Elite XL von Bayer Diagnostics Fernwald [26-28] und das etwas größere Tischgerät HemoCue B-Glucose Analyzer von HemoCue GmbH Großostheim [29-34] gewählt. Diese fünf POCT-Glukometer werden im fortlaufenden Text mit Accutrend, Accu-Chek1, Accu-Chek2, Elite, HemoCue abgekürzt. Die Messverfahren dieser POCT-Glukometer beruhen auf dem enzymatisch amperometrischen Prinzip, wobei das Enzym entweder Glukoseoxidase (Elite, Accu-Chek) oder Glukosedehydrogenase (HemoCue, Accutrend) ist [22]. Bei der Umsetzung von Glukose durch das auf einem Träger immobilisierte Enzym entstehen Elektronen, die auf einen Akzeptor übertragen werden, der als Mediator zwischen dem Enzym und der Graphitableitung für den Elektronenfluss durch das Messgerät sorgt.

POCT-Glukometer zur Glukose-Selbstbestimmung benötigen nur minimale Mengen an Probenmaterial (ein Tropfen kapilläres Vollblut aus der Fingerbeere), die auf Einwegteststreifen bzw. Mikrokuvetten appliziert dem Mess-System zugeführt werden, und bestimmen innerhalb von Sekunden die Glukosekonzentration. Die Geräte sind vorrangig für die Messung im Kapillarblut ausgerichtet, können aber auch arterielle und venöse Blutproben messen. Zwei Accu-Chek Glukometer derselben Bauart wurden in den Methodenvergleich aufgenommen, so dass ein limitierter Vergleich gleicher Systeme möglich war.

Als nicht portables POCT-Gerät wurde das Blutgasanalysegerät AVL Omni Modular von Roche Diagnostics GmbH Mannheim (im fortlaufenden Text mit Omni abgekürzt) mit in den Methodenvergleich aufgenommen. Es ist für den intensivmedizinischen und ambulanten Bereich konzipiert und misst neben der Glukose weitere Notfallmessgrößen, wie z.B. die Blutgasanalyse, Laktat und Elektrolyte. Die Analyse der Glukosekonzentration erfolgt beim Omni in der wässrigen Phase über eine direkte Glukose-Oxidase-Methode, wobei die Glukoseaktivität gemessen wird. Die wässrige Phase enthält die Konzentration, die am Wirkort zur Verfügung steht, entspricht somit also der physiologisch wirksamen Konzentration („Glukoseaktivitäts“-Messung). Diese Methode wird als unbeeinträchtigt durch Hämatokrit- und Protein-Interferenz erachtet. Das Messen ist aufwendiger als mit den portablen POCT-Geräten. Die zu untersuchende Probenart wird aus einer 20µl-Kapillare in das Gerät „eingesaugt“ und dann mechanisch durch ein Schlauchsystem in die Messkammer

transportiert und danach entsorgt. Das Gerät bedarf einer regelmäßigen Wartung mit Wechsel von Flüssigkeitsreagenzien, Leerung von Abfallbehältern, etc.

Als weiteres Routineanalysegerät wurde Cobas Integra 700 von Roche Diagnostics GmbH Mannheim [35-39] (im fortlaufenden Text mit Integra abgekürzt), das seit 1994 auf dem Markt ist, in den Methodenvergleich mit aufgenommen. Integra war wie Ebio ein Analysegerät des Zentrallabors des Klinikums Bremen-Mitte und erfüllte demzufolge entsprechende Qualitätskriterien. Es bestimmt die Glukosekonzentration mit der Hexokinase/Glukose-6-Phosphat-Dehydrogenase-Methode über ein amperometrisch-photometrisches Prinzip [22;40;41]. Glukose wird in Anwesenheit von Hexokinase mit ATP zu Glukose-6-Phosphat phosphoryliert, welches unter katalytischer Wirkung der Glukose-6-Phosphat-Dehydrogenase (G-6-PDH) mit NADP⁺ zu 6-Phosphogluconat und NADPH₂ reagiert. Die NADPH₂-Zunahme wird bis zum Stillstand der Reaktion gemessen, die ermittelte Extinktionszunahme ist proportional der Glukosekonzentration im Testansatz.



Das Einbeziehen eines zweiten etablierten Laboranalysegerätes in den Methodenvergleich erfolgte, um den Diskordanzraten-Test (siehe Kapitel 3.3.3) auf seine Validität zu überprüfen. Unter der Annahme, dass die beiden Standardgeräte aufgrund der gleichen Messgenauigkeit die gleichen Messwerte liefern würden, konnte somit die Verlässlichkeit des Diskordanzraten-Tests in der Erkennung der Messunterschiede der POCT-Geräte überprüft werden.

Während des Studienzeitraums (siehe Kapitel 3.2.2.4) wurden die portablen POCT-Geräte ausschließlich für den Studienzweck genutzt. Es wurden keine Glukometer zwischenzeitlich ausgetauscht. Angaben zu Gerätespezifikationen, Gerätezubehör und verwendeten Chargen sind in Tabelle 3 erfasst. Da nicht immer alle Geräte eingesetzt wurden, ergeben sich unterschiedliche „Einsatzfrequenzen“ und Anwendungszeiträume der verschiedenen Geräte, die in Tabelle 4 dargelegt werden.

Tabelle 3: Gerätespezifikationen, Chargen der Teststreifen

Gerätename	Hersteller (Ort)	Messprinzip (Enzym)	Messsystem (weiteres Zubehör)	benötigtes Probenvolumen	Chargen der Teststreifen
EBIO plus 6668	Eppendorf AG (D-22331 Hamburg)	Glukoseoxidase enthaltende Elektrode (enzymatisch – amperometrisch)	2 ml Safe-Lock-Gefäße 500 µl ESAT-Hämolyseflösung 10 µl End-zu-End-Kapillaren	10 µl	
Cobas Integra 700	Roche Diagnostics GmbH (D-68305 Mannheim)	HK / G6P-DH ¹⁾ (Flüssig-Reagentien) (enzymatisch – photometrisch)	Gerätespezifische Verschlusskappen	10 µl	
AVL Omni Modular	Roche Diagnostics GmbH (D-68305 Mannheim)	MSS-Sensoren (Flüssig-Reagentien) (colorimetrisch)	200 µl End-zu-End-Kapillare	60 µl	
HemoCue B-Glukose Analyzer	HemoCue GmbH (D-63762 Grossostheim)	Glukosedehydrogenase Absorptionsspektrometer (enzymatisch – photometrisch)	Mikroküvetten (HemoCue B-Glukose Küvetten)	5 µl	70495, 80545, 110830, 110808, 1010004
Glucometer Elite XL	Bayer Diagnostics (D-35463 Fernwald)	Glukoseoxidase, Kaliumhexacyanoferrat(III) (enzymatisch – amperometrisch)	Teststreifen (Glukometer Elite Sensor)	2 µl	Charge F5, Lotnummern: GM05SC, OA05HM, E05EP, OH05JS, OH05OS, OJ05DM, OJ05NS, GJ05RM
Accutrend Sensor	Roche Diagnostics GmbH (D-68305 Mannheim)	Glukosedehydrogenase Hexacyanoferrat(III) (enzymatisch – amperometrisch)	Teststreifen (Accutrend Sensor Comfort Glukose)	4 µl	433864, 433931, 433923, 434142, 317, 346
Accu-Chek Plus	Roche Diagnostics GmbH (D-68305 Mannheim)	Glukoseoxidoreduktase (enzymatisch – amperometrisch)	Teststreifen (Accu-Chek Sensor Comfort Glukose)	2 µl	Accu-Chek 1: 228310-30, 228310-80, 350 Accu-Chek 2: 228217-80, 216, 346

¹⁾ G6P-DH = Glukose-6-Phosphat-Dehydrogenase, HK = Hexokinase

Tabelle 4: Anwendungszeiträume und Einsatzfrequenz der Glukometer während des Studienzeitraums

Gerätename	Anwendungszeitraum	n ²⁾
Ebio plus 6668	15.06.99 bis 07.09.01 (26 Monate + 3 Wochen)	78
Cobas Integra 700	15.06.99 bis 07.09.01 (26 Monate + 3 Wochen)	46
AVL Omni Modular	03.10.99 bis 07.09.01 (23 Monate + 1 Woche)	70
HemoCue B-Glucose	02.08.00 bis 25.05.01 (9 Monate + 3 Wochen)	44
Glukometer Elite XL	04.09.00 bis 09.07.01 (10 Monate + 1 Woche)	47
Accutrend Sensor	18.07.00 bis 10.07.01 (11 Monate + 3 Wochen)	53
Accu-Chek Plus 1.Gerät	25.09.00 bis 04.07.01 (9 Monate + 1 Woche)	40
Accu-Chek Plus 2.Gerät	20.07.00 bis 13.07.01 (11 Monate + 3 Wochen)	44

¹⁾ Studienkollektivzeitraum = 26 Monate + 3 Wochen

Beginn des Studienkollektiv-Zeitraums (erstes Vergleichsgerät im Einsatz): 15.06.99

Ende des Studienkollektiv-Zeitraums (letztes Vergleichsgerät im Einsatz): 07.09.01

²⁾ Einsatzfrequenz der Glukometer

(entsprechend der Anzahl der Versuchstage, der Anzahl der Patienten und der Anzahl der oGTT)

3.1.2 Materialien und Reagenzien

Die verwendeten Materialien und Reagenzien sind in Tabelle 5 und 6 aufgelistet.

Tabelle 5: Materialien

Material	Handels-Name	Hersteller	Bestell-Nr.
Stechhilfe	Softclix II Stechhilfe	Roche Diagnostics GmbH, D-68305 Mannheim	PZN: 1 834 670 Art-Nr. 10114
Lanzetten	Softclix II Lanzetten	Roche Diagnostics GmbH, D-68305 Mannheim	18021
Kapillarhalter	Kapillarhalter für die End-to-End-Kapillare	Sarstedt AG, D-51582 Nümbrecht	95.1048
200 µl-Kapillare	heparinisierte 200 µl End-zu-End-Kapillaren	Roche Diagnostics GmbH, D-68305 Mannheim	MC 0024; 3113477180
10 µl-Kapillare, Mikrokapillare (Ebio)	End-to-End-Kapillare, 10 µl, NH ₄ -Heparin	Sarstedt AG, D-51582 Nümbrecht	19.414
Glukosevorlage (Ebio)	Mikro-Probengefäß mit anhängendem PE-Softstopfen, 500 µl ESAT-Hämolyse-lösung (Safe-Lock-Gefäße)	Sarstedt AG, D-51582 Nümbrecht	41.1443.902
Braunüle	Venflon 2, 2HG	Ohmeda AB, S-25106 Helsingborg	1452-2
Mandrin	Safelon Obturator	Ohmeda AB, S-25106 Helsingborg	4252-3
Monovette	S-Monovette, 2,7 ml, Lithium-Heparin	Sarstedt AG, D-51582 Nümbrecht	05.1553
Integra-spezifische Verschlusskappe	Reagiergefäß, 0,5 ml mit Verschlusskappe für Cobas Integra	Roche Diagnostics GmbH, D-68305 Mannheim	2056925001, US#1044869
Eppendorfgefäß	Reagiergefäß, 1,5 ml mit anhängendem EASY CAP, Polypropylen, farblos	Sarstedt AG, D-51582 Nümbrecht	72.690.550

Tabelle 6: Reagenzien

Substanz / Reagenz	Hersteller	Bestell- / Chargen-Nr.
Dextro O.G-T	Roche Diagnostics GmbH, D-68305 Mannheim	PZN: 7 145 670
D(+)-Glucose (wasserfrei)	National Institute of Standards and Technology Gaithersburg, MD 20899, USA	SRM 917a
D(+)-Glucose (wasserfrei)	Merck, D-64271 Darmstadt	8337; K24170237 743
Validate A	Organon Teknika GmbH, D-69214 Eppelheim	Lot.No.6B403
Validate N	Organon Teknika GmbH, D-69214 Eppelheim	Lot.No.6B401
Human-Albumin (pur+lyophilisiert)	Serva GmbH, D-69115 Heidelberg	11870; Lot.28060
Natrium-Chlorid	Merck, D-64271 Darmstadt	6404; K26811304 952

3.2 METHODEN

3.2.1 Standard- und Kontroll-Lösungen

Eine wichtige Maßnahme, um die Messqualität von Analysegeräten zu überprüfen, ist das regelmäßige Messen von Standard- und Kontroll-Lösungen, bei denen die Konzentration der Messgröße bekannt ist. Die Analysegeräte werden so hinsichtlich ihrer Präzision und Richtigkeit überprüft. Die Hersteller geben meist sowohl ein bestimmtes Zeitintervall als auch die Anzahl an Messungen an, innerhalb dessen bzw. nach denen Qualitätskontrollen durchgeführt werden sollen, wobei das Ereignis, welches jeweils als erstes eintritt, den Kontrollzeitpunkt entscheidet. POCT-Glukometer werden in der Regel mit hypo- und hyperglykämischen und manchmal auch noch zusätzlich mit normoglykämischen Kalibrier/Kontroll-Lösungen geliefert, sogenannte spezifische geräteabhängige Lösungen. Geräteunabhängige handelsfertige Kontroll-Lösungen enthalten mehrere Messgrößen (Proteine, Elektrolyte, Spurenelemente, etc.), die analysiert werden können. Tägliches Messen von Standard- und Kontroll-Lösungen ist als Qualitätssicherungsmaßnahme integrativer Bestandteil des Laborroutinebetriebes, und notwendig, um tägliche geräteimmanente Messschwankungen von Analysegeräten zu erkennen, und das Ausmaß der von-Tag-zu-Tag-Präzision zu ermitteln.

Anhand eines täglich aus den Messungen der Standard- oder Kontroll-Lösungen abgeleiteten Quotienten der Messabweichung gegenüber der eingewogenen Konzentration der Messgröße können tägliche Messschwankungen kalibriert bzw. korrigiert werden („Tages-bezogene Korrektur“). Entscheidend dabei ist, dass die für eine solche Korrektur verwendete Standard- oder Kontroll-Lösung im Hinblick auf diesen Zweck in der Anwendung des zu korrigierenden Analysegerätes überprüft wurde. Zu beachten ist, dass sich nicht jedes Analysegerät gleichermaßen durch eine solche Tages-bezogene Korrektur kalibrieren lässt bzw. dass sich nicht jede Standard- oder Kontroll-Lösung gleichermaßen für die Tages-bezogene Korrektur der Messergebnisse verschiedener Analysegeräte eignet.

In dieser Studie wurden Messungen mit vier Standard- und Kontroll-Lösungen zur Qualitätssicherung und zur von-Tag-zu-Tag-Präzision durchgeführt. Die Standardlösungen (Glu-NIST, Glu-Alb) wurden selbst angesetzt, als Kontroll-Lösungen wurden zwei handelsfertige lyophilisierte verwendet, Val A und Val N (Tabelle 6 in Kapitel 3.1.2, Tabelle 7). Aus Gründen der Vergleichbarkeit wurden anstelle der von den Firmen vorgesehenen Kalibrier/Kontroll-Lösungen diese geräteunabhängigen Kontroll-Lösungen eingesetzt.

Tabelle 7: Glukosestandards und Kontroll-Lösungen

	Glu-NIST	Glu-Alb	Val A	Val N
Hersteller	NIST ¹⁾	Merck / Serva	Organon Teknika	Organon Teknika
Charge	SRM 917a	K24170237 743 / 28060	6B403	6B401
Sollwert ²⁾ [mg/dl]	200	200	207,95	93,88
Vertrauensbereich ³⁾ [mg/dl]			176,78 – 239,04	79,82 – 107,94
15%-Bereich [mg/dl]	170 – 230	170 – 230	176,8 – 239, 1	79,8 – 108,0

1) NIST = National Institute of Standards and Technology

2) Sollwert gemäß den Angaben des Herstellers bzw. eingewogene Glukosekonzentration des eigenen Ansatzes

3) Vertrauensbereich gemäß den Angaben des Herstellers

Glu-Nist war eine reine selbst angesetzte wässrige Glukoselösung (wasserfreie D(+)-Glukose) in destilliertem Wasser, welche zusätzlich noch 900 mg/dl NaCl enthält. Als Glukose wurde die des National Institute of Standards and Technology (NIST) verwendet, die durch den höchsten zertifizierten Reinheitsgrad von 99,9 ausgezeichnet war. Glu-Alb war ein selbst angesetzter proteinhaltiger wässriger Glukosestandard (D(+)-Glukose der Firma Merck), der zusätzlich noch 8 g/l humanes Albumin und 900 mg/dl NaCl enthält. Beide Glukosestandards (Glu-NIST und Glu-Alb) waren auf 200 mg/dl Glukose eingewogen; NaCl wurde den Glukosestandards zugefügt, da Omni nur in elektrolyt-angereicherten Lösungen messen kann.

Val A und Val N (Validate A und Validate N) waren handelsfertige proteinhaltige Kontroll-Lösungen (lyophilisierte Präzisions- und Richtigkeitskontrollseren) der Firma Organon Teknika. Sie lagen als gefriergetrocknete Produkte vor und wurden mit destilliertem Wasser entsprechend den Anweisungen des Herstellers angesetzt. Val A enthielt Glukose in hoher Konzentration von 207,95 mg/dl, Val N hingegen in niedriger Konzentration von 93,88 mg/dl. Die vom Hersteller angegebenen Vertrauensbereiche, die sich aus 3 Standardabweichungen oberhalb und unterhalb des eingewogenen Wertes (Sollwert) ergeben, lagen beim Val A zwischen 176,76 und 239,04 mg/dl und beim Val N zwischen 79,80 und 107,96 mg/dl.

Da der Vertrauensbereich der lyophilisierten Kontroll-Lösungen 15% um den Sollwert lag, wurde für die wässrigen Glukoselösungen ein „15%-Bereich“ als Vertrauensbereich angegeben (15% ober- und unterhalb der eingewogenen 200 mg/dl Glukosekonzentration), der zwischen 170 und 230 mg/dl lag.

Die Standardlösungen Glu-NIST und Glu-Alb wurden während des gesamten Studienzeitraumes in einem mit einem Gummipropfen luftdicht verschlossenen Erlenmeyerkolben im Kühlschrank bei 8° Celsius gelagert. Wegen Kontaminationsgefahr bei wiederholtem Gebrauch und nicht-tiefgefrorener Lagerung wurden sie in Abständen von maximal 3 Monaten immer wieder neu angesetzt. Die Validate-Lösungen wurden in kleinen Portionen von 100 µl in Eppendorfgefäße abgefüllt und bei -22° Celsius gelagert. Pro Versuchstag wurde jeweils der Inhalt eines Eppendorfgefäßes aufgebraucht.

3.2.1.1 Mess-Serien, Sollwert-Abweichung, Tages-bezogene Korrektur

Eine Mess-Serie bezieht sich auf alle Messungen einer Lösung über den gesamten Studienzeitraum. An jedem Versuchstag während des Studienzeitraumes wurden die Glukosekonzentrationen der beiden Standard- und beiden Kontroll-Lösungen mit jedem Glukometer doppelt bestimmt, die Doppeltbestimmungen wurden gemittelt („tägliche Mittelwerte“). Somit ergaben sich pro Gerät 4 Mess-Serien (Glu-NIST-Serie, Glu-Alb-Serie, Val A-Serie, Val N- Serie).

Die täglichen Mittelwerte der Mess-Serien wurden arithmetisch gemittelt und mit dem jeweiligen Sollwert verglichen. Aus den prozentualen Mittelwert-Abweichungen vom Sollwert („Sollwert-Abweichung“) wurde die Messgenauigkeit der Geräte abgeleitet. Messgenau galten Geräte, deren Sollwert-Abweichung 5% nicht überschritt, was in Anlehnung an die

Richtlinie der Bundesärztekammer {304} als Akzeptanzbereich definiert wurde. Die von-Tag-zu-Tag-Präzision wurde pro Gerät für jede der 4 Mess-Serien anhand der Berechnung des Variationskoeffizienten (VK%) bestimmt.

Eine Tages-bezogene Korrektur erfolgte für jedes Gerät durch den täglich aus den Messungen des Glu-NIST-Standard abgeleiteten Quotienten der Messabweichung gegenüber der eingewogenen 200 mg/dl Glukosekonzentration. Tägliche Messschwankungen wurden mit diesem Quotienten, dem „Tages-bezogenen Korrekturfaktor“, korrigiert. Messwerte wurden als „Tages-korrigiert“ bezeichnet, nachdem die unkorrigierten Messwerte mit dem Tages-bezogenen Korrekturfaktor multipliziert worden waren. Messwerte wurden daraufhin hinsichtlich ihres „Korrekturstatus“ in unkorrigiert (uk) oder Tages-korrigiert (korr) differenziert.

3.2.2 Biologisches Probenmaterial

Im Rahmen von medizinisch indizierten oralen Glukosetoleranztests wurden Glukosekonzentrationen in drei verschiedenen Probenarten (Kapillarblut, Venenblut und venöses Plasma) bestimmt, um somit Messwerte zu erhalten, die eine diagnostische Aussagekraft bezüglich eines Diabetes mellitus oder einer Glukoseintoleranz haben.

3.2.2.1 Oraler Glukosetoleranztest

Ein oraler Glukosetoleranztest (oGTT) ist indiziert, wenn es darum geht, die Verdachtsdiagnose eines Diabetes mellitus bei grenzwertiger Nüchternglukosekonzentration abzusichern. Die in dieser Studie durchgeführten oGTT erfolgten gemäß den Empfehlungen der Weltgesundheitsorganisation [42]. In Konsens mit der Ethikkommission Bremen wurden die Probanden in dieser Studie mindestens drei Tage vor dem oGTT vom Stationsarzt und in schriftlicher Form über den Ablauf der Untersuchung und die vorbereitenden Maßnahmen aufgeklärt. Das Einverständnis an der Teilnahme musste durch Unterschrift des Probanden gegeben werden.

Um nüchtern zum oGTT zu erscheinen, der morgens zwischen 8 und 9 Uhr begonnen wurde, musste mindestens acht Stunden vor Beginn des Testes Nahrungskarenz eingehalten werden. Als Glukosebelastungstrunk wurde Dextro O.G-T. von Roche

Diagnostics GmbH genommen, der als wirksamen Bestandteil ein Mono- und Oligosaccharid-Gemisch (Stärkehydrolysat) enthält, das nach enzymatischer Spaltung einer Menge von 75 g wasserfreier Glukose entspricht. Die Einnahme des Trunks erfolgte zügig innerhalb von fünf Minuten. Blutentnahmen erfolgten in halbstündlichen Intervallen zu fünf vordefinierten Zeitpunkten: im Nüchternzustand, sowie 30, 60, 90 und 120 Minuten nach der Glukosebelastung. Während des Testes mussten sich die Probanden körperlich ruhig (sitzend oder liegend) verhalten. Abgesehen von der acht-stündigen Nahrungskarenz wurden keine diätetischen Restriktionen verlangt; die Probanden sollten vielmehr in den Tagen vor dem Test normale Ernährungsgewohnheiten einhalten und sich kohlenhydratreich ernähren. Übermäßige körperliche Aktivität sollte außerdem in den Tagen vor dem Test vermieden werden. Pro Versuchstag wurde ein oGTT an einem Probanden durchgeführt.

3.2.2.2 Gewinnung des Probenmaterials und Analyse der Glukosekonzentration

Pro Entnahmezeitpunkt des oGTT wurde kapilläres und venöses Vollblut parallel abgenommen; Plasma wurde aus dem venösen Vollblut innerhalb von maximal 10 Minuten abzentrifugiert. Die Glukosekonzentration wurde in allen drei Probenarten (Kapillarblut = CB, venöses Vollblut = VB und venöses Plasma = VP) doppelt bestimmt, deren Werte täglich gemittelt wurden („tägliche Mittelwerte“). Die genauen Angaben zu den verwendeten Reagenzien, Materialien, Gerätezubehör und der in den POCT-Glukometern verwendeten Sensoren sind in den Tabellen 3, 5 und 6 (in den Kapiteln 3.1.1.2 und 3.1.2) beschrieben.

Kapilläre Blutproben wurden von der Fingerbeere entnommen, nachdem der erste Tropfen nach der Punktion mittels Softclix II-Stechhilfe verworfen wurde: Für die Ebio-Messungen wurden zwei NH₄-heparinisierte 10 µl End-zu-End-Kapillaren luftblasenfrei und vollständig gefüllt, außen abgewischt, sofort je in ein Mikroprobengefäß mit 500 µl ESAT-Hämolyselösung transferiert und kräftig geschüttelt. Danach wurden für die POCT-Glukometermessungen zwei heparinisierte 200 µl End-zu-End-Kapillaren zu etwa 60% gefüllt. Die Venenblutentnahme wurde mit Lithium-heparinhaltigen 2,7 ml Monovetten über eine venöse Verweilkanüle unmittelbar nach der Kapillarblutentnahme durchgeführt.

In einem Nachbarraum, getrennt von dem Raum, in dem die Blutentnahmen erfolgten, wurden die Analysen durchgeführt. Mit den ersten Tropfen der mit Kapillarblut gefüllten 200 µl End-zu-End-Kapillaren wurden die Sensoren bzw. Küvetten von Accutrend, Accu-Chek, HemoCue und Elite beschickt, der Rest wurde in das Omni-Gerät eingeführt. Aus den mit Venenblut gefüllten Monovetten wurden analog zwei heparinisierte 10 µl End-zu-End-

Kapillaren mit Venenblut gefüllt und in die Glukosevorlagen mit 500 µl hämolysierendes Reagenz gegeben und zwei heparinisierte 200 µl Kapillaren bis zu 60% mit Venenblut gefüllt, mit denen dann die Sensoren der POCT-Geräte beschickt wurden. Das restliche Venenblut wurde zentrifugiert (spätestens 10 Minuten nach der Blutentnahme). Die Glukosekonzentrationsbestimmungen des venösen Plasmas erfolgten analog den Vollblutmessungen. Aus dem venösen Plasma wurde neben der Glukosekonzentration auch noch Insulin und der HbA1c bestimmt, um anhand eines pathologischen Insulinsekretionsmusters und eines erhöhten HbA1c die Diagnose eines Diabetes mellitus bzw. einer gestörten Glukosetoleranz abzusichern [43-45].

Die mit den POCT-Geräten durchgeführten Messungen erfolgten innerhalb von 15 Minuten nach jeder Blutentnahme. Die mit den Routineanalysegeräten (Ebio und Integra) durchgeführten Messungen erfolgten direkt im Anschluss an den oGTT zuerst mit dem Ebio und anschließend aus den gleichen Hämolysatansätzen mit dem Integra, wobei die Hämolysatansätze aus den Ebio-spezifischen Mikro-Probengefäßen in Integra-spezifische Gefäße überführt werden mussten. Glukosekonzentrationen waren in hämolysierten Proben mindestens 24 Stunden stabil.

Alle Messungen wurden, um Handhabungsunterschiede zu minimieren, unter standardisierten Bedingungen im Zentrallaboratorium durch die Autorin der hier vorliegenden Arbeit durchgeführt, wobei die Anweisungen der Hersteller zur Bedienung der Geräte (Raumtemperatur, Verweilzeit der Proben auf den Teststreifen etc.) eingehalten wurden. Die Glukosekonzentrationen wurden, soweit genügend Probenmaterial gewonnen wurde, immer doppelt bestimmt. Das Kapillarblutmaterial reichte nicht immer für Doppelbestimmungen aus, je nachdem wie schnell die Punktionsstelle an der Fingerkuppe versiegte. Eine zweite Kapillarpunktion wurde mit Rücksicht auf die Probanden vermieden. Die Doppelbestimmung der Glukosekonzentration durch Ebio und Integra erfolgte immer, da die kapilläre Blutentnahme immer mit der Befüllung der beiden 10 µl End-zu-End-Glaskapillaren begonnen wurde, bevor die 200 µl End-zu-End-Kapillaren für die POCT-Bestimmungen gefüllt wurden.

Verwertet wurden in der vorliegenden Studie nur die Nüchtern- und die zwei-Stunden Postprandial-Messungen.

3.2.2.3 Mess-Serien und Vergleiche der verschiedenen Probenarten

Eine Mess-Serie entspricht den Messungen in einer Probenart mit einem Gerät zu einem Entnahmezeitpunkt und Korrekturstatus. Aufgrund dreier verschiedener Probenarten (Kapillarblut, venöses Vollblut, venöses Plasma), zweier Entnahmezeitpunkte (nüchtern, 2h-postprandial) und zweier Korrekturstatus (uk, korr) ergeben sich somit pro Gerät 12 Mess-Serien.

Die Mess-Serien der Testmethode werden mit den entsprechenden – d.h. in Probenart, Entnahmezeitpunkt und Korrekturstatus übereinstimmenden - Mess-Serien der Referenzmethode verglichen. Es wird zwischen „intrasystemischen“ und „intersystemischen“ Vergleichen unterschieden.

Intrasystemische Vergleiche entsprechen den Vergleichen der Erst- mit den Zweitmessungen eines Analysegerätes. Die intrasystemischen Vergleiche wurden mit den unkorrigierten Messwerten berechnet, also den Messwerten, die vom Display der Glukometer direkt abgelesen bzw. vom Analysegerät ausgedruckt wurden. Aufgrund der Anzahl der Analyseverfahren ($n = 8$), der Probenarten ($n = 3$), der Messzeitpunkte ($n = 2$) und des unkorrigierten Korrekturstatus ($n = 1$) gab es infolgedessen 48 intrasystemische Vergleiche. Anhand der intrasystemischen Vergleiche wurde die Messgenauigkeit der Testgeräte in sich beurteilt, die Reproduzierbarkeit der Messergebnisse überprüft, und die Präzision in der Serie analysiert, außerdem wurden die intrasystemischen Diskordanzraten berechnet (Kapitel 3.3.3).

Intersystemische Vergleiche entsprachen den Vergleichen der gemittelten Doppelmessungen der Testgeräte mit den parallel bestimmten gemittelten Doppelmessungen des Referenzgerätes Ebio. Es wurden sowohl die unkorrigierten als auch die Tages-korrigierten Testgerät-Messwerten mit den Tages-korrigierten Ebio-Messwerten verglichen, die bezüglich Probenart und Messzeitpunkt übereinstimmten. Aufgrund der Anzahl der Testverfahren ($n = 7$), der Probenart ($n = 3$), des Messzeitpunktes ($n = 2$) und des Korrekturstatus ($n = 2$) wurden 84 intersystemische Vergleiche durchgeführt. Anhand der intersystemischen Vergleiche wurde die Messgenauigkeit der Testgeräte im Vergleich zum Routineanalysegerät (Ebio) beurteilt und die diagnostische Leistungsfähigkeit und diagnostische Fehlerrate bezüglich einer Entscheidungsgrenze, die intersystemische Diskordanzrate (Kapitel 3.3.3), abgeschätzt. Es wurden darüber hinaus noch 42 transformierte intersystemische Vergleiche vorgenommen, was in Kapitel 3.3.2 näher erläutert wird, um zu überprüfen, wie sich die Elimination des Bias auf die Messgenauigkeit der Testgerät auswirkt.

3.2.2.4 Probanden, Kollektive und Zeiträume

Folgende Begriffe „Studienprobanden“, „Studienzeitraum“, „Pool der Studienprobanden“, „Populationsprobanden“, „Populationskollektiv“, „Pool der Populationsprobanden“, „Populationszeitraum“, „Studienkollektive“ werden differenziert und im nachfolgenden Text definiert.

Die Probanden, die in diese Studie eingeschlossen wurden, die sogenannten „Studienprobanden“, waren erwachsene Personen, an denen prospektiv von Juni 1999 bis September 2001 im Zentrallabor des Klinikums Bremen-Mitte orale Glukosetoleranztests zur Abklärung einer diabetischen Stoffwechsellage durchgeführt wurden. Die Glukosebestimmungen erfolgten bei den Studienprobanden sowohl mit dem Ebio-Analysegerät als auch mit mindestens einem der Testgeräte. Die Studienprobanden kamen vorwiegend aus zwei Tageskliniken (Innere Medizin und Hautklinik) des Klinikums Bremen-Mitte. Es waren Personen, die aufgrund ihrer Familienanamnese oder anderer prädisponierender Risikofaktoren gefährdet waren, ein metabolisches Syndrom und somit eine pathologische Glukosetoleranz zu entwickeln [46], oder Personen, die an chronisch-dermatologischen Erkrankungen litten, die möglicherweise durch einen noch nicht diagnostizierten Diabetes mellitus ausgelöst waren. Aus der Studie ausgeschlossen waren Minderjährige, schwerkranke Patienten und solche, bei denen ein Diabetes mellitus bekannt war. Zu der vorliegenden Studie liegt unter dem Antrag „Studie zur Früherkennung des Typ 2 Diabetes mellitus“ ein Votum der Ethikkommission der Ärztekammer des Landes Bremens mit der Nummer 67 vor.

Der Zeitraum, in dem die Studienprobanden rekrutiert wurden, wurde als „Studienzeitraum“ bezeichnet. Da nicht immer Messungen mit allen Glukometern durchgeführt wurden, unterscheiden sich „Anwendungszeiträume“ und „Einsatzfrequenz“ der einzelnen Geräte (siehe Tabelle 4 im Kapitel 3.1.1.2).

Innerhalb des Studienzeitraums wurden 78 Studienprobanden rekrutiert. Diese 78 Studienprobanden entsprechen dem „Pool der Studienprobanden“. Anhand der mit dem Ebio-Analysegerät gemessenen Glukosekonzentrationen des venösen Plasmas zum Nüchtern- und 2h-postprandial-Zeitpunkt, der HbA1c-Werte und der Insulinsekretionsmuster wiesen gemäß den WHO-Vorgaben [42;47-49] 16 Personen eine eindeutig diabetische, 34 Personen eine nicht-diabetische (bzw. glukosetolerante) und 25 Personen eine intermediäre (bzw. glukose-intolerante) Stoffwechsellage auf; 3 Personen wurden aufgrund von Besonderheiten in ihrer Anamnese oder aufgrund unplausibler Glukosemuster auf keine

Diagnose festgelegt. Die von der WHO erstellten Diagnosekriterien sind in Tabelle 8 im Anhang zusammengefasst.

Neben den Studienprobanden wurden in dieser Studie auch noch sogenannte „Populationsprobanden“ für die statistische Auswertung mittels des Diskordanzraten-Testes (siehe Kapitel 3.3.3) rekrutiert. Zur Diskordanzraten-Berechnung ist nämlich ein repräsentatives „Populationskollektiv“ notwendig, welches sich in diesem Falle aus Personen zusammensetzt, bei denen die Verdachtsdiagnose eines Diabetes mellitus abgeklärt werden soll. Alle Personen, bei denen im Zentrallabor des Klinikums Bremen-Mitte seit Einführung des Ebio-Analysegerätes 1994 klinisch indizierte oGTTs durchgeführt worden waren, wurden in den „Pool der Populationsprobanden“ eingeschlossen. Daten vor dem Studienzeitraum waren handschriftlich aufgezeichnet worden, und wurden somit retrospektiv erhoben; Daten während und nach dem Studienzeitraum hingegen wurden prospektiv erhoben. Der Pool der Studienprobanden entsprach einer Untermenge des Pools der Populationsprobanden. Während des „Populationszeitraums“, der sich über 7½ Jahre erstreckte, wurden 168 Personen in den Pool der Populationsprobanden eingeschlossen.

Gemäß den WHO-Vorgaben [42;47-49] (bezogen auf die Ebio-Messwerte des venösen Plasmas, Insulinsekretionsmuster und HbA1c) wiesen 39 Personen eine diabetische, 75 Personen eine nicht-diabetische und 46 Personen eine intermediäre Stoffwechsellage auf. Acht Personen blieben unklassifiziert, weil die Daten unvollständig oder inkonsistent waren. In den Tabellen 9 und 10 sind die Anzahl der Messungen in den verschiedenen Probenarten und die Patientencharakteristika (Geschlecht, BMI, Diagnose) der Populationsprobanden zusammengefasst.

Tabelle 9: Anzahl an Messungen der Populationsprobanden pro Probenart

Probenart	nüchtern	2h-postprandial
Kapillarblut	150	168
Venenblut	167	146
Venöses Plasma	160	156

Tabelle 10: Patientencharakteristika der Populationsprobanden

Glukose-Stoffwechsellage	Sex	Anzahl			BMI-Range	ohne BMI-Angabe
		insgesamt	<= 25 BMI	>= 25 BMI (30 BMI)		
nicht-diabetisch	alle	75*	39	32 (20)	16 bis 63	4*
	m	33	15	17 (5)	16 bis 63	1
	f	41	15	24 (15)	17 bis 49	2
intermediär	alle	46	12	33 (21)	22 bis 53	1
	m	25	5	19 (11)	22 bis 42	1
	f	21	7	14 (10)	22 bis 53	0
diabetisch	alle	39	10	29 (14)	21 bis 59	0
	m	18	5	13 (8)	21 bis 59	0
	f	21	5	16 (6)	21 bis 55	0
Sondergruppen	alle	8	3	5 (2)	19 bis 41	0
	m	3	1	2 (0)	19 bis 41	0
	f	5	2	3 (2)	23 bis 27	0
alle**	alle	168	64 (2 < 19 BMI)	99 (26)	16 bis 63	5
	m	79				
	f	89				

Abkürzungen: BMI = Body Mass Index, m = maskulin, f = feminin

* bei einem keine Angabe zum Geschlecht ** Mittelwert des BMI: 29 +/-8; Median des BMI: 27

Für die Berechnung der Diskordanzraten werden für jeden Methodenvergleich zwei Kollektive benötigt: ein „Studienkollektiv“ und ein repräsentatives „Populationskollektiv“. Aus dem Pool der Studienprobanden wird das jeweilige Studienkollektiv und aus dem Pool der Populationsprobanden das entsprechende Populationskollektiv rekrutiert. Ein Kollektiv war in dieser Studie jeweils durch Probenart (Kapillarblut, venöses Vollblut, venöses Plasma), Messzeitpunkt (nüchtern, postprandial) und Korrekturstatus (uk, korr) charakterisiert. Für die intrasystemischen Vergleiche (Kapitel 3.2.2.3) aller 8 Glukometer (unkorrigierte Messwerte), wurden infolgedessen 48 Studienkollektive gebildet. Für die intersystemischen Vergleiche (Kapitel 3.2.2.3) der 7 Testverfahren (unkorrigierte und Tages-korrigierte Messwerte) gegenüber den Tages-korrigierten Ebio-Messwerten wurden 84 Studienkollektive ermittelt. Da die vergleichende Komponente beim Populationskollektiv entfällt und nur die Tages-korrigierten Ebio-Messwerte als Referenzwerte genommen wurden, beschränkte sich die Anzahl an Populationskollektiven insgesamt auf n = 6.

3.3 STATISTISCHE AUSWERTUNG

Zunächst wurden die Raten an falsch positiven und falsch negativen Messergebnissen bezüglich einer Entscheidungsgrenze durch reines Auszählen berechnet, um einen prima vista Eindruck der Verteilung der Werte bzw. der Rate an falschen Zuordnungen zu erhalten. Die statistische Auswertung der Messwerte erfolgte dann durch konventionelle Methodenvergleiche. Die Ergebnisse daraus stellten die Grundlage für die Auswertung des Diskordanzraten-Tests (DR-Test) dar, ein neuer biometrischer Ansatz, der für „entscheidungsorientierte“ Methodenvergleiche entwickelt wurde (Kapitel 3.3.3).

3.3.1 Ermittlung der Rate falsch positiver und falsch negativer

Messwerte

Die diagnostische Zuverlässigkeit wird klassischerweise durch Ermittlung falsch und richtig positiver und negativer Werte bezüglich einer Entscheidungsgrenze ermittelt, woraus dann die Gütekriterien eines diagnostischen Testes (z.B. Sensitivität und Spezifität) abgeleitet werden. Die Effizienz eines diagnostischen Testes berechnet sich durch die Rate an falsch positiven und falsch negativen Messwerten im Vergleich zur Gesamtzahl der Wertepaare hinsichtlich einer Entscheidungsgrenze und wird als Prozentzahl angegeben (Tabelle 11 im Anhang).

Alle 48 intrasystemischen und 84 intersystemischen Vergleiche (siehe Kapitel 3.2.2.3) wurden mithilfe von Microsoft Excel „ausgezählt“ und die Rate an falsch positiven und falsch negativen Messwerten berechnet. Als Entscheidungsgrenzen wurden die WHO-definierten Grenzen zugrundegelegt, mit denen ein Diabetes mellitus diagnostiziert wird (Tabelle 8 im Anhang), die also zwischen „nicht-diabetisch“ und „diabetisch“ bzw. zwischen „glukoseintolerant“ und „diabetisch“ unterscheiden [42;47-49]. Je nach Probenart und Messzeitpunkt wurden demzufolge folgende Entscheidungsgrenzen für die Glukosekonzentrationen zugrundegelegt. Eine diabetische Stoffwechsellage liegt vor, wenn:

	Kapillarblut	Venöses Blut	Venöses Plasma
Nüchtern [mg/dl]	≥ 110	≥ 110	≥ 126
oder 2h-postprandial [mg/dl]	≥ 200	≥ 180	≥ 200

3.3.2 Konventioneller Methodenvergleich

Alle intrasystemischen und intersystemischen Vergleiche wurden auf eine lineare Beziehung zwischen Ebio und den Testmethoden mittels linearer Regressionsanalysen geprüft. Da die klassische lineare Regressionsanalyse davon ausgeht, dass die Referenzmethode nur „wahre“ Daten liefert, die aber bei Messung biologischer Matrizes nicht vorliegen können, wurden die Regressionsanalysen mit dem Strukturmodell nach Passing/Bablok durchgeführt, weil die lineare Regressionsanalyse nach Passing/Bablok ausdrücklich Fehler in beiden Methoden zulässt [50;51].

Berechnet wurde die Funktion der Ausgleichsgeraden $Y = \alpha + \beta \cdot X + \varepsilon$ (α = Achsenabschnitt, β = Steigung, ε = Zufallsfehler bzw. Residuum) mit Angabe des 95%-Konfidenzintervalls und Signifikanzniveaus, außerdem folgende Größen: arithmetischer Mittelwert, Median, Minimum, Maximum und Standardabweichung der X- und Y-Reihen, der Korrelationskoeffizient r und die Streuung der Residuen md_{95} . Bei Passing/Bablok entspricht md_{95} der Standardabweichung der Residuen ($s_{y,x}$), die sonst bei der klassischen linearen Regressionsanalyse angegeben wird.

Die Messgenauigkeit der Testgeräte gegenüber dem Referenzgerät Ebio wurde bewertet, indem die prozentuale Abweichung der gemittelten Messwerte der Testmethode (*Y-quer*) mit den gemittelten Messwerten der Referenzmethode (*X-quer*) ermittelt wurde. Die Mittelwert-Abweichung wurde berechnet über die Funktion: $f(X) = (Y - X) \cdot 100 / X$. Übereinstimmung der Messwerte zweier Analysegeräte wurde angenommen, wenn deren mittlere Konzentrationen weniger als 5% voneinander abwichen.

Bei der Interpretation der Ergebnisse der linearen Regression muss beachtet werden, dass Störeinflüsse, wie z.B. zu kleine Spannweite des Konzentrationsbereichs (Maximum/Minimum-Ratio < 2,0), Auftreten von Ausreißern (Messwerte außerhalb des vierfachen Standardfehlers der Residuen), Vorliegen von Heteroskedasizität einen linearen bzw. nicht-linearen Zusammenhang verfälschen kann. Durch graphische Präsentation der Messwerte, beispielsweise in X/Y-Streudiagrammen oder Differenzenplots (z.B. nach Bland/Altman [1]), können solche Störeinflüsse visualisiert werden. Die Linearität wurde in dieser Studie visuell und durch den Cusum-Test geprüft. Störgrößen und Ausreißer wurden im X/Y-Diagramm (Abbildung 3), in dem Ausgleichsgerade und Winkelhalbierende ($y = x$) eingezeichnet wurden, und durch Differenzenplot (Abbildung 4) detektiert.

Abbildung 3: X/Y-Diagramm bei linearer Regression

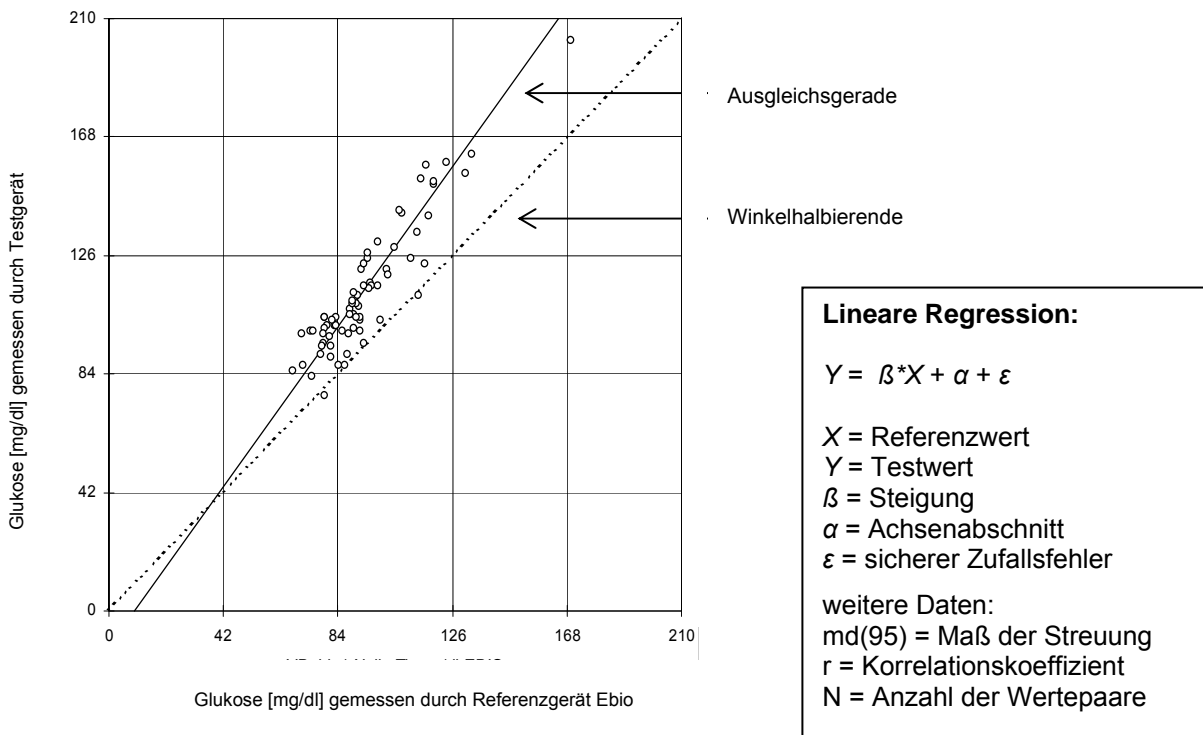
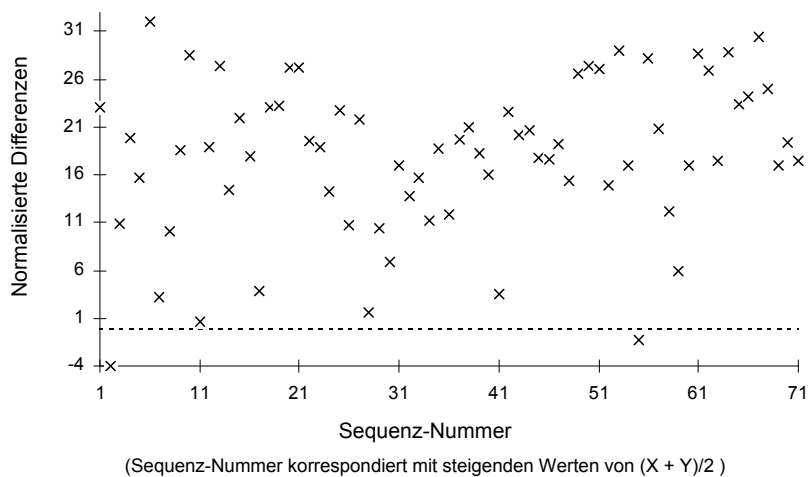


Abbildung 4: Differenzenplot normalisierter Differenzen



Der Differenzenplot ist eine Sonderform eines Punktediagramms, in dem die Differenzen beider Methoden gegen den Mittelwert der beiden Methoden aufgetragen werden. Dadurch kann man optisch beurteilen, wie hoch die Schwankungsbreite der Übereinstimmung ist, ob eine Messmethode prinzipiell höher oder tiefer misst als die andere (systematischer

Messfehler), und ob die Abweichung der Methoden oder die Streuung der Abweichungen von der Höhe der Werte abhängig ist. Dies ist zwar auch aus dem X/Y-Diagramm der linearen Regression ablesbar, aber für Ungeübte ist es deutlicher aus dem Differenzenplot erkennbar. Die Sequenz-Nummer (Abszisse) korrespondiert mit den steigenden Werten aus den Mittelwerten der Summe der Wertepaare aus X und Y: $(X + Y) / 2$. Orientierend ist der Wert, in dem keine Differenz vorliegt, als gestrichelte Nulllinie eingezeichnet.

Die Berechnungen erfolgten mit einem Computer-unterstützten Evaluationsprogramm von Barembruch *et al.* [52] mithilfe eines für Microsoft-Excel 98 optimierten Add-On-Programm: EVAPAK for Windows© Version 3.1 (Evaluation Program for Method Comparison von Boehringer Mannheim GmbH).

Trotz linearer Beziehung können die Werte der Testmethode von den „wahren“ Referenzwerten abweichen, was einer systematischen Verzerrung entspricht und als Bias bezeichnet wird. Der Bias kann konstant oder proportional zu X oder einer Kombination von beiden sein. Der Bias ist visuell erkennbar und kann durch die Regressionsanalyse geschätzt und von den beobachteten Testwerten subtrahiert werden, wodurch die Kalibrationsgerade sozusagen auf die Winkelhalbierende verschoben wird. Diese Bias-Korrektur ist nichts anderes als eine nachträgliche Kalibration, die hier als Transformation bezeichnet wurde und auf alle intersystemischen Vergleiche angewendet wurde. Dabei ist anzumerken, dass die Streuung der Messwerte durch die Transformation unbeeinflusst bleibt.

3.3.3 Diskordanzraten-Test

Der Diskordanzraten-Test ermöglicht dem Anwender, die Rate an diagnostischen Fehlentscheidungen bei Methodenvergleichen in Bezug auf eine Entscheidungsgrenze unter Berücksichtigung der Verteilung der untersuchten Messgröße im Patienten- bzw. Probandenkollektiv abzuschätzen. Er ist ein neuer biometrischer Ansatz, der von R. Haeckel und W. Wosniok während des Studienzeitraumes entwickelt wurde, um entscheidungsorientierte Methodenvergleiche durchzuführen [8-14]. In dieser Studie wurde der Diskordanzraten-Test erstmalig angewendet.

Die Entwicklung des Diskordanzraten-Testes erfolgte vor dem Hintergrund, dass die Messung einer Variable an ein und derselben Probe mit verschiedenen Methoden in der Regel zu unterschiedlichen Resultaten führt. Dieser Unterschied wird für die Praxis relevant, wenn weitere Entscheidungen davon abhängen, ob das Messergebnis über oder unter einer Entscheidungsgrenze liegt. Die Absicht des entscheidungsorientierten Methodenvergleichs besteht darin, die Wahrscheinlichkeit divergenter Entscheidungen abzuschätzen.

Mit dem Diskordanzraten-Test wird die diagnostische Messgenauigkeit der Methoden aufgrund von diskordanten Messwerten bzw. Diskordanzen bezüglich der Entscheidungsgrenze, die Erkrankte von Nicht-Erkrankten unterscheidet, beurteilt. Konkordante Messungen liegen vor, wenn die gepaarten Messungen beider Analyseverfahren entweder unterhalb oder oberhalb der Entscheidungsgrenze liegen. Diskordante Messungen liegen vor, wenn die Messwerte der Testmethoden „auf der anderen Seite“ der Messwerte der Referenzmethode liegen. Eine positive Diskordanz bedeutet, dass der mittels Testmethode erhobene Messwert oberhalb der Entscheidungsgrenze liegt, während der des Referenzverfahrens unterhalb der Entscheidungsgrenze liegt. Entsprechend bedeutet eine negative Diskordanz, dass der mittels Testmethode erhobene Messwert unterhalb der Entscheidungsgrenze liegt, während der des Referenzverfahrens oberhalb der Entscheidungsgrenze liegt. Positive und negative Diskordanzen können mit falsch positiven und falsch negativen Messungen gleichgesetzt werden. Die positiven und negativen Fehlerraten (Diskordanzraten = DR) hingegen, die aus den Diskordanzen berechnet werden, und mit denen letztlich die diagnostische Messgenauigkeit der Testmethoden bewertet wird, entsprechen nur annäherungsweise der Rate an falsch positiven oder falsch negativen Messungen, da die Diskordanzraten-Berechnung auf einem stochastischen Berechnungsansatz beruht.

3.3.3.1 Erforderliche Daten und Berechnungsschritte des Diskordanzraten-Tests

Für die Anwendung des Diskordanzraten-Tests bedarf es folgender Daten:

- a) die Messwerte eines Methodenvergleiches (Mess-Serie der Testmethode und gepaarte Mess-Serie der Referenzmethode), was als Studienkollektiv bezeichnet wird (siehe Kapitel 3.2.2.4).
- b) eine repräsentative Mess-Serie derselben Messgröße, die als Populationskollektiv bezeichnet wird.

Das Populationskollektiv entspricht einem repräsentativen Kollektiv, in dem der zu bewertende diagnostische Test normalerweise durchgeführt wird. Idealerweise sollten die Messwerte mit derselben Referenzmethode bestimmt worden sein, mit dem konsekutiv auch der entscheidungsorientierte Methodenvergleich durchgeführt wurde. Es können aber auch Messwerte, die in einem anderen Labor bestimmt wurden, sofern der Bias bekannt und vernachlässigbar ist, für das Populationskollektiv herangezogen werden. Um eine typische Verteilung eines repräsentativen Kollektivs zu erhalten, wird eine Mindestanzahl an Populationsdaten gefordert ($n > 100$), aus denen die Verteilung abgeleitet wird.

- c) Entscheidungsgrenzen der beiden Methoden, die miteinander verglichen werden.

Die Berechnung des Diskordanzraten-Tests erfolgt in 5 Schritten:

- (1) Lineare Regression: Vergleich der beiden Methoden z.B. mittels dem linearen Strukturmodell nach Passing-Bablok [50].
- (2) „Analytische“ Wahrscheinlichkeit: Bestimmung der Wahrscheinlichkeit, dass die Testmethode für einen als gegeben angenommenen Referenzwert einen divergenten Wert liefert.
- (3) „Populationsbasierte“ Wahrscheinlichkeit: Bestimmung der Wahrscheinlichkeit, dass ein Messwert in dem zu untersuchenden Kollektiv überhaupt vorkommen kann.

- (4) „Verknüpfte Wahrscheinlichkeit“: Verknüpfung der analytischen Wahrscheinlichkeit mit der populationsbasierten Wahrscheinlichkeit, dass der in (2) als gegeben angenommene Referenzwert in einer repräsentativen Population auftritt.
- (5) „Zusammenfassende“ Wahrscheinlichkeit (Diskordanzrate): Summation der verknüpften Wahrscheinlichkeiten zur zusammenfassenden Wahrscheinlichkeit:
- a. aller mit der Referenzmethode bestimmten Werte, die unterhalb der Entscheidungsgrenze liegen zur Gesamtwahrscheinlichkeit, dass eine Diskordanz in Form von falsch positiven Entscheidungen auftritt (= positive DR, DR(pos)), sowie
 - b. aller mit der Referenzmethode bestimmten Werte, die oberhalb der Entscheidungsgrenze liegen zur Gesamtwahrscheinlichkeit, dass eine Diskordanz in Form von falsch negativen Entscheidungen auftritt (= negative DR, DR(neg)).
 - c. Die Summe aus DR(pos) und DR(neg) ergibt dann die Gesamt-DR (DR(ges)).

Schritt (1) Im ersten Schritt wird mittels linearer Regression (z.B. klassische lineare Regression oder lineares Strukturmodell nach Passing/Bablok) untersucht, welche Beziehung zwischen Referenzmethode und Testmethode besteht. Resultat dieses Schrittes ist sowohl eine Ausgleichsgerade für die Beziehung zwischen beiden Methoden wie auch die Kenntnis der Streuung (Standardabweichung) um diese Gerade. Eine systematische Verzerrung wird durch die Abweichung der Ausgleichsgeraden von der Winkelhalbierenden ersichtlich (Ausmaß der Steigungs-Abweichung von 1 und / oder der Achsenabschnitts-Abweichung von 0). Das Ausmaß der Unsicherheit bei der Bestimmung der Geraden wird durch den Zufallsfehler ε wiedergegeben und ist zusammen mit dessen Standardabweichung $s_{y,x}$ bzw. md_{95} Ausdruck der Streuung um die Gerade.

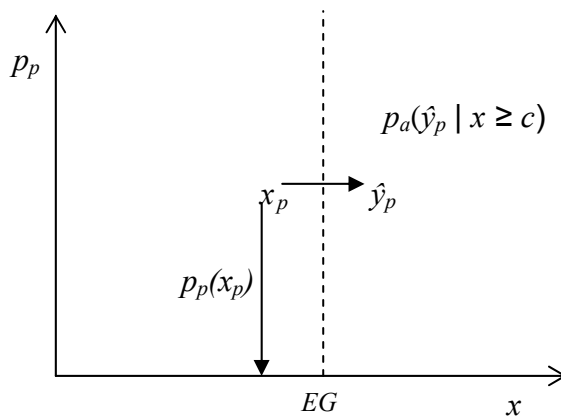
Schritt (2) Als nächster Schritt wird die „analytische“ Wahrscheinlichkeit bestimmt, die aussagt, mit welcher Wahrscheinlichkeit die Testmethode für dieselbe Probe einen diskordanten Wert liefert, mit welcher Wahrscheinlichkeit also der Testwert (Y) auf der „anderen Seite“ der Entscheidungsgrenze des korrespondierenden Referenzwertes (X) liegt.

Dafür wird systematisch jeder mögliche Messwert der Referenzmethode daraufhin untersucht, welcher Wert für dieselbe Probe seitens der Testmethode zu erwarten ist. Dazu wird die Beziehung, die in Schritt (1) bei der linearen Regression bestimmt wurde,

herangezogen. Dieses Wissen wird benötigt, um zu berechnen, mit welcher Wahrscheinlichkeit ein Testwert durch das Zusammenwirken von systematischer Verzerrung und Zufallsfehler über die Entscheidungsgrenze gerät, obwohl der parallel mit der Referenzmethode gemessene Messwert unter der Entscheidungsgrenze liegt, und es somit zu einer diskordanten Entscheidung in Form einer falsch positiven Entscheidung kommt, bzw. analog dazu, mit welcher Wahrscheinlichkeit ein Testwert durch das Zusammenwirken von systematischer Verzerrung und Zufallsfehler unter die Entscheidungsgrenze gerät, obwohl der parallel mit der Referenzmethode gemessene Messwert auf oder über der Entscheidungsgrenze liegt, und es somit zu einer diskordanten Entscheidung in Form einer falsch negativen Entscheidung kommt.

Die vollständigen Formeln zur Berechnung der analytischen Wahrscheinlichkeiten wurden 2003 publiziert [12]. Das Prinzip des Diskordanzraten-Testes wird in Abbildung 5 graphisch dargestellt mit Nennung der Kernformeln; in den Abbildungen 6 und 7 (Kapitel 3.3.3.2) werden die analytischen Wahrscheinlichkeiten positiver und negativer Diskordanzen graphisch verdeutlicht.

Abbildung 5: Prinzip des Diskordanzraten-Tests



<p>falsch positiv: $P(x_p < c, \hat{y}_{p,\varepsilon} \geq c) = \sum_{x_p < c} p_p(x_p) \cdot p(\hat{y}_{p,\varepsilon} \geq c \mid x_p)$</p> <p>falsch negativ: $P(x_p \geq c, \hat{y}_{p,\varepsilon} < c) = \sum_{x_p \geq c} p_p(x_p) \cdot [1 - p(\hat{y}_{p,\varepsilon} \geq c \mid x_p)]$</p>

p_p = Populationswahrscheinlichkeit
 p_a = analytische Wahrscheinlichkeit
 x = X-wert
 x_p = gemessener X-Wert innerhalb der Population
 y_p = gemessener Y-Wert innerhalb der Population
 $y_{p,\varepsilon}$ = y_p unter Berücksichtigung des Zufallsfehler
 EG = Entscheidungsgrenze
 ε = Zufallsfehler in Methodenvergleichen

Die Wahrscheinlichkeit solcher Verschiebungen über die Entscheidungsgrenze hinweg ist nicht konstant, sondern um so größer, je näher der erwartete Testwert an der Entscheidungsgrenze liegt, und je größer die beteiligte Streuung ist.

Die analytische Wahrscheinlichkeit quantifiziert also das Fehler-Risiko, das beim Vergleich zweier Methoden erwartet wird. Die Schätzung der analytischen Wahrscheinlichkeit beim Diskordanzraten-Test ermöglicht, dass aus einem viel kleineren als sonst in Methodenvergleichen benötigten Datenset aussagekräftige Ergebnisse berechnet werden können.

Schritt (3) Die Betrachtungen in den Schritten (1) und (2) berücksichtigen nicht, dass in der Praxis Proben aus einem realen Kollektiv zu untersuchen sind, dessen Messwerte keineswegs eine Gleichverteilung auf der Skala der möglichen Werte aufweisen, sondern vielmehr eine für das Kollektiv spezifische Verteilung. Häufen sich Werte in der Nähe der Entscheidungsgrenze, so ist die analytische Wahrscheinlichkeit von diskordanten Entscheidungen erhöht. Der Diskordanzraten-Test muss deswegen für eine Gesamtaussage über das Verhalten der beiden Verfahren bei Anwendung auf das fragliche Kollektiv die nach Erfahrung zu erwartende Verteilung der Messwerte im Kollektiv berücksichtigen. Zu diesem Zweck wird auf der Grundlage der Messwerte eines repräsentativen Populationskollektivs (gemessen mit der Referenzmethode) die „populationsbasierte“ Wahrscheinlichkeit des Auftretens eines jeden möglichen Messwertes mittels numerischer Integration bestimmt. Den Berechnungen liegt eine Gauss Kernel Dichteschätzung zugrunde [53], die durch Kreuzvalidation geglättet wird [54]. Fluktuationen, die bei auf empirischer Häufigkeit basierenden Berechnungen in kleinen Stichproben auftreten, sind bei solcher Vorgehensweise ausgeglichen. Graphisch resultiert eine Dichteschätzungskurve, die Populationskurve (siehe Abbildung 8 in Kapitel 3.3.3.2).

Schritt (4) Die Wahrscheinlichkeit, dass ein möglicher Referenzmesswert im Kollektiv auftritt und bei Verwendung der Testmethode diskordant eingestuft wird, ergibt sich dann aus Verknüpfung der zuvor bestimmten einzelnen Wahrscheinlichkeiten (analytischer und populationsbasierter Wahrscheinlichkeit), getrennt nach den Situationen Referenzwert x unterhalb der Entscheidungsgrenze bzw. Referenzwert x auf oder oberhalb der Entscheidungsgrenze. Diese „verknüpfte“ Wahrscheinlichkeit wird für jeden Referenzwert einzeln berechnet.

Schritt (5) Tatsächlich ist aber nicht eine Aussage für einen einzelnen bestimmten Referenzwert x gesucht, sondern eine zusammenfassende Aussage über alle x , die vorkommen können. Diese „zusammenfassende“ Wahrscheinlichkeit, dass innerhalb des gesamten Kollektivs diskordante Entscheidungen erfolgen, ergibt sich durch Summation der verknüpften Wahrscheinlichkeiten, wiederum getrennt nach den beiden Arten von diskordanten Entscheidungen.

Aus diesem letzten Schritt resultieren die positiven und negativen Diskordanzraten. Die positive DR entspricht der Gesamtwahrscheinlichkeit für alle unterhalb der Entscheidungsgrenze liegenden Referenzwerte, dass die mit der Testmethode gemessenen Werte auf oder oberhalb der Entscheidungsgrenze liegen, was zu einer falsch positiven Entscheidung führt. Die negative DR entspricht der Gesamtwahrscheinlichkeit für alle auf oder oberhalb der Entscheidungsgrenze liegenden Referenzwerte, dass die mit der Testmethode gemessenen Werte unterhalb der Entscheidungsgrenze liegen, was zu einer falsch negativen Entscheidung führt. Die Summe aus positiver und negativer DR ergibt dann die Gesamt-DR.

Anzumerken ist, dass die „Wahrscheinlichkeit für diskordante Entscheidung“ dasselbe ist wie der (zu erwartende) Anteil bzw. die (zu erwartende) Rate von diskordanten Entscheidungen an allen Entscheidungen. Die Wahrscheinlichkeiten für diskordante (falsch positive bzw. falsch negative) Entscheidungen sind summarische Größen für den entscheidungsorientierten Methodenvergleich, der sowohl die Beziehungen zwischen den beiden Methoden einschließlich der darin enthaltenen Unsicherheit, die jeweilige Entscheidungsgrenze und die Verteilung der untersuchten Größe im Kollektiv berücksichtigt.

Die Gesamt-DR entspricht der Wahrscheinlichkeit, mit der divergente Entscheidungen gefällt werden, die positive DR entspricht der Wahrscheinlichkeit, mit der falsch positive Entscheidungen gefällt werden, die negative DR entspricht der Wahrscheinlichkeit, mit der falsch negative Entscheidungen gefällt werden.

3.3.3.2 Graphische schematische Darstellung des Diskordanzraten-Tests

Abbildung 6: Beziehung zwischen Referenzmesswert x (nach oben weisender Pfeil) und erwartetem Testwert \hat{y} (nach links weisender Pfeil) und der analytischen Wahrscheinlichkeit (schraffierte Fläche), dass der reale Testwert oberhalb der Entscheidungsgrenze (EG) = 110 mg/dl liegt, obwohl der Referenzwert ($x = 75$ mg/dl) unter der EG liegt. Die Fehlerverteilung rund um \hat{y} ist übertrieben groß gezeichnet. Die schraffierte Fläche ist $p(y \geq c | x)$ (Freihandzeichnung!) und beschreibt die analytische Wahrscheinlichkeit, dass eine Messung mit der Testmethode durch Zusammenwirken von systematischer Verzerrung mit Zufallseffekten über die EG gerät. Weiße Fläche unterhalb der gezeichneten Kurve = Wahrscheinlichkeit, dass keine Diskordanz vorliegt.

Abbildung 6: Analytische Wahrscheinlichkeiten positiver Diskordanzen

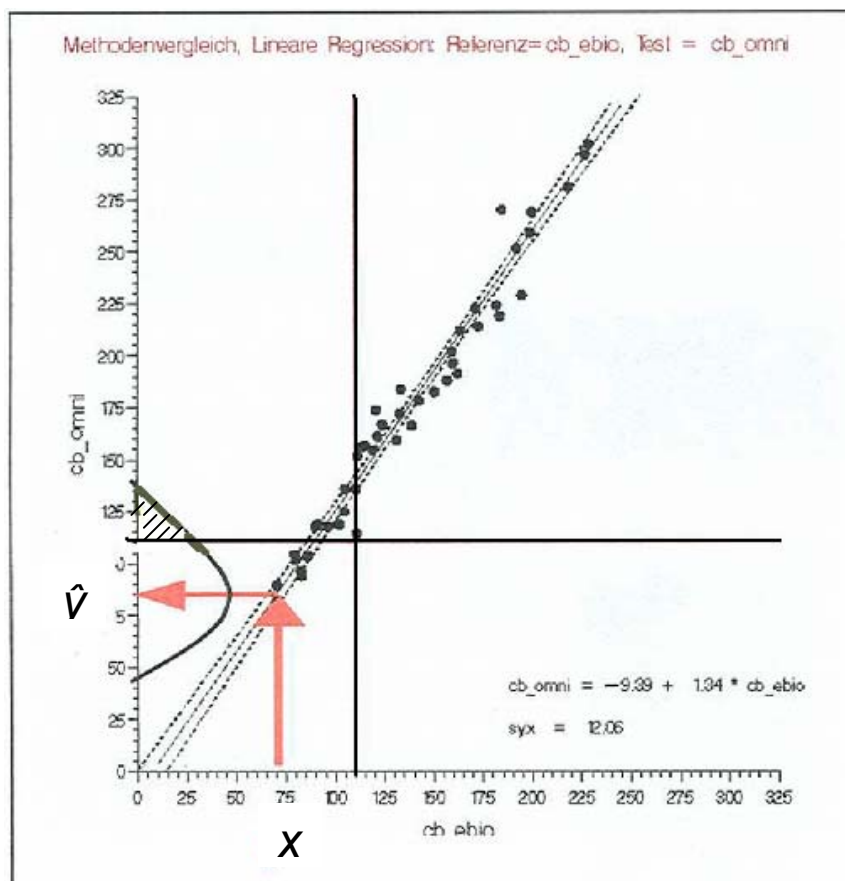


Abbildung 7: Beziehung zwischen Referenzmesswert x (nach oben weisender Pfeil) und erwartetem Testwert \hat{y} (nach links weisender Pfeil) und der analytischen Wahrscheinlichkeit (schraffierte Fläche), dass der reale Testwert unterhalb der EG = 110 mg/dl liegt, obwohl der Referenzwert ($x = 125$ mg/dl) über der EG liegt. Die Fehlerverteilung rund um \hat{y} ist übertrieben groß gezeichnet. Die schraffierte Fläche ist $p(y < c | x) = 1 - p(y \geq c | x)$ (Freihandzeichnung!) und beschreibt die analytischen Wahrscheinlichkeit, dass eine Messung mit der Testmethode durch Zusammenwirken von systematischer Verzerrung mit Zufallseffekten über die EG gerät. Weiße Fläche oberhalb der gezeichneten Kurve = Wahrscheinlichkeit, dass keine Diskordanz vorliegt.

Abbildung 7: Analytische Wahrscheinlichkeiten negativer Diskordanzen

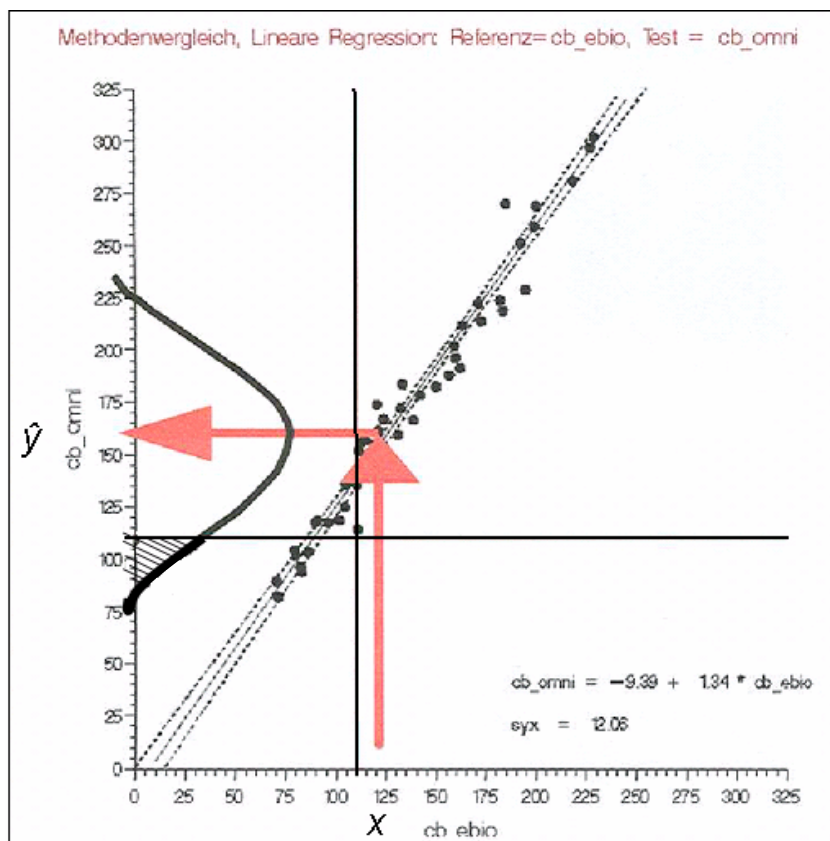
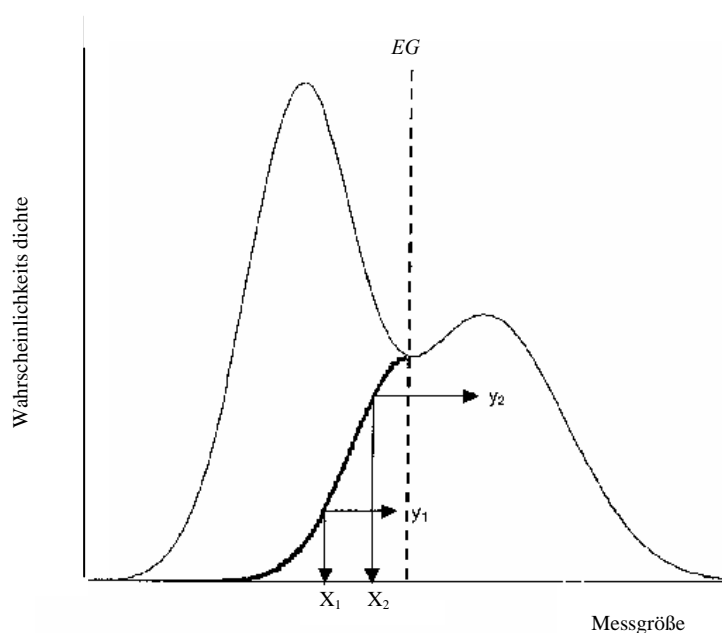


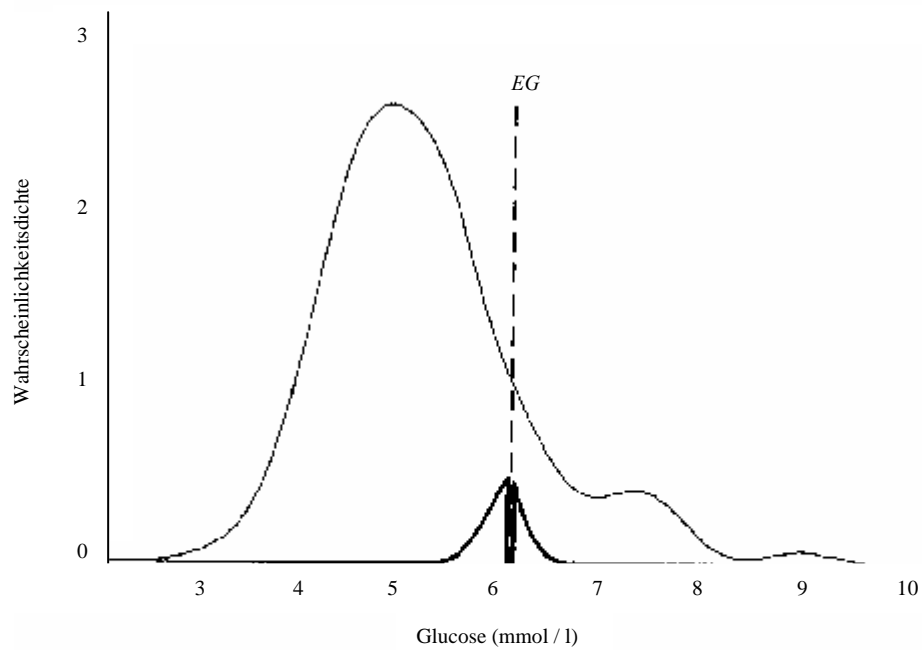
Abbildung 8: Schematische Präsentation der Wahrscheinlichkeiten. Die horizontale Achse bezieht sich auf die Größe der zu bestimmenden Messgröße x (in dieser Arbeit mg/dl); die vertikale Achse bezieht sich auf die Wahrscheinlichkeitsdichte der Messwerte. Die obere Kurve (Dichteschätzungskurve oder Populationskollektivkurve) repräsentiert die Wahrscheinlichkeit $p_{\text{pop}}(x)$, dass x als Messwert in der repräsentativen Population, in der der diagnostische Test durchgeführt wird, auftritt. Eigentlich könnte dies durch jeweils einen senkrechten Strich für jedes ganzzahlige x dargestellt werden, jedoch würde dann das Bild zu dicht. So sind nur die oberen Endpunkte dieser Striche verbunden und ergeben den Eindruck einer durchgezogenen glatten Kurve. Die Fläche unterhalb der unteren Kurve links von der Entscheidungsgrenze EG beschreibt die Gesamtwahrscheinlichkeit des Auftretens von positiven Diskordanzen des gepaarten Vergleiches beim Methodenvergleich (die Summe aller analytischen Wahrscheinlichkeiten $p(y \geq EG | x)$), die Fläche unterhalb der unteren Kurve rechts von der Entscheidungsgrenze (hier nicht eingezeichnet) beschreibt die Wahrscheinlichkeit des Auftretens von negativen Diskordanzen des gepaarten Vergleiches beim Methodenvergleich. Die Fläche unterhalb der unteren Kurve kann als prozentualer Anteil zur Fläche unterhalb der Dichteschätzungskurve ausgedrückt werden. Dieser Prozentsatz entspricht der DR. Die DR hängt ab von der Verteilung der gepaarten Messwerte um die Entscheidungsgrenze, vom Bias zwischen Test- und Referenzmethode und insbesondere dem zufälligen Fehler um die Regressionsgerade, der die Streuung angibt. Kurz zusammengefasst: Obere Kurve = repräsentative Population. EG = Entscheidungsgrenze. x_2 (Referenzmethode) = nicht-erkrankt. y_2 (Testmethode) = erkrankt. Positive Diskordanzen = Fläche unter der Diskordanzkurve links von c (hier Fläche unter der unteren Kurve). Negative Diskordanzen = Fläche unter der Diskordanzkurve rechts von EG (hier nicht eingezeichnet).

Abbildung 8: Dichteschätzung und positive Diskordanzkurve, theoretisches Darstellung



In Abbildung 9 wird anhand eines praktischen Beispiels die Verteilung des Populationskollektiv und der Diskordanzraten um die Entscheidungsgrenze *EG* für Glukosekonzentration im Nüchtern-Kapillarblut dargestellt. Die Vergleichs- und Populationswerte wurden durch das Ebio-Analysegerät bestimmt, die Testwerte mit dem Elite-Glukometer.

Abbildung 9: Dichteschätzung und Diskordanzkurven, praktisches Beispiel (Nüchtern-Kapillarblut, Vergleich zwischen Elite- und Ebio-gemessenen Glukosekonzentrationen)



3.3.3.3 Anwendung des Diskordanzraten-Tests in der vorliegenden Studie

Für alle 48 intrasystemischen, 84 intersystemischen und 42 transformierten intersystemischen Vergleiche (Kapitel 3.2.2.3, Kapitel 3.3.2) wurden die Diskordanzraten bezüglich der von der WHO festgelegten diabetischen Entscheidungsgrenzen (Kapitel 3.3.1, Tabelle 8 im Anhang [42;47-49]) berechnet. Der Messwert-Datensatz eines Vergleiches entsprach dem jeweiligen Studienkollektiv. An repräsentativen Populationskollektiven gab es $n = 6$, die sich in Probenart und Zeitpunkt unterschieden (siehe Kapitel 3.2.2.4).

Glukometer, deren Gesamt-DR einen Wert von 5,0% nicht überschritt, wurden als „gleichwertig“ wie die Referenzmethode angesehen. Glukometer, deren Gesamt-DR einen Wert von 10% nicht überschritt und deren positive und negative DR maximal 5,0% groß waren, wurden als „noch akzeptabel“ erachtet. Glukometer, deren Gesamt-DR über 10% oder deren positive oder negative DR über 5,0% lagen, galten für diagnostische Zwecke als „nicht akzeptabel“.

Die DR-Berechnungen erfolgten mit einem von R. Haeckel und W. Wosniok neu entwickelten Berechnungsprogramm, welches ursprünglich mit dem SAS STAT Paket und der SAS Programmiersprache (SAS Institute) programmiert worden war, und dann in Microsoft-Excel 2000 umgesetzt wurde. Der Diskordanzraten-Test war so programmiert, dass die einzelnen Berechnungsschritte im Hintergrund abliefen (Lineare Regression, Berechnungen der analytischen Wahrscheinlichkeiten für falsch positive und falsch negative Diskordanzen, Berechnungen der Populationskollektiv-Verteilungen, Berechnungen der populationsbasierten Wahrscheinlichkeiten). Als Ergebnis lieferte der Diskordanzraten-Test die positiven und negativen Diskordanzraten und die korrespondierenden „Konkordanzraten“.

4 ERGEBNISSE

4.1 MESSUNGEN DER STANDARD- UND KONTROLL-LÖSUNGEN

4.1.1 Sollwert-Abweichungen und Impräzision

Tabelle 12 enthält die Ergebnisse die aus den Mess-Serien der Standard- und Kontroll-Lösungen aller Glukometer berechnet wurden (Kapitel 3.2.1.1): Anzahl der Messungen, Mittelwert, Standardabweichung, Minimum und Maximum, Sollwert-Abweichung und Variationskoeffizient. Die Mittelwerte werden zusammen mit ihren Standardabweichungen in den Abbildungen 10 a – d graphisch präsentiert.

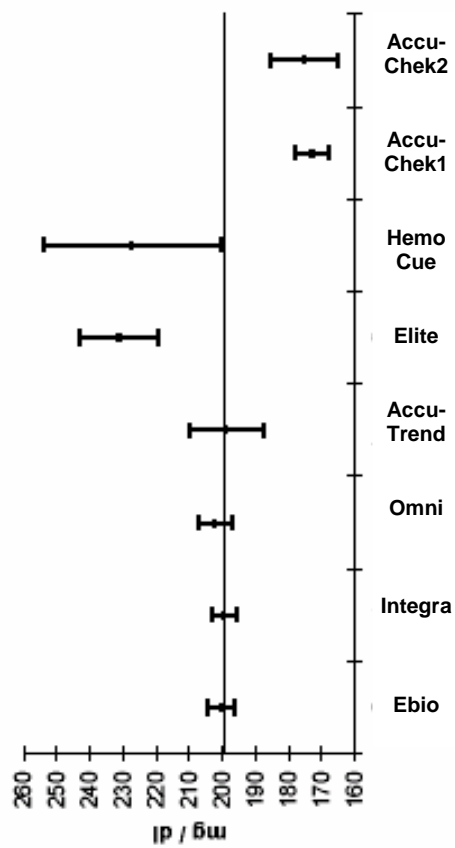
Tabelle 12: Messergebnisse in den Standard- und Kontroll-Lösungen

Lösung (SW)		Ebio	Integra	Omni	Accu-trend	Elite	Hemo Cue	Accu-Chek1	Accu-Chek2
Glu-Nist (200)	MW	200,4	199,5	202,2	198,9	231,2	227,4	173,0	175,3
	SD	3,8	3,8	4,9	11,2	11,9	26,7	5,2	10,1
	Min	188,0	182,0	182,1	169,5	171,5	169,3	110,0	142,0
	Max	217,8	215,7	215,0	230,0	272,3	277,5	217,5	192,5
	VK%	1,9	1,9	2,4	5,7	5,2	12,2	3,1	2,4
	△%	0,2	-0,3	1,1	-0,5	15,6	13,7	-13,5	-12,3
	N	82	41	65	56	49	44	40	47
Glu-Alb (200)	MW	199,1	197,1	204,3	193,7	242,7	305,0	209,5	210,5
	SD	9,9	12,8	16,2	11,9	11,0	27,4	12,3	10,8
	VK%	4,9	6,5	7,9	6,1	4,5	8,4	5,9	5,1
	△%	-0,5	-1,4	2,2	-3,2	21,4	52,5	4,8	5,2
	N	44	37	35	40	35	30	30	34
Val A (208)	MW	205,9	201,4	211,2	204,7	225,8	215,1	163,4	164,1
	SD	8,1	6,8	8,6	10,8	15,9	6,5	5,4	5,6
	VK%	4,0	3,4	4,1	5,3	7,4	3,1	3,3	3,4
	△%	-1,0	-3,2	1,5	-1,6	8,5	2,0	-21,4	-21,1
	N	68	40	59	51	47	42	38	42
Val N (93,9)	MW	96,0	91,4	97,1	88,6	103,2	115,3	82,7	86,8
	SD	5,4	4,9	4,0	6,1	4,7	5,2	3,8	4,1
	VK%	5,6	5,3	4,0	7,1	4,6	4,6	4,5	4,8
	△%	2,3	-2,7	3,4	-6,7	10,0	22,8	-11,9	-7,5
	N	63	37	54	47	42	41	37	40

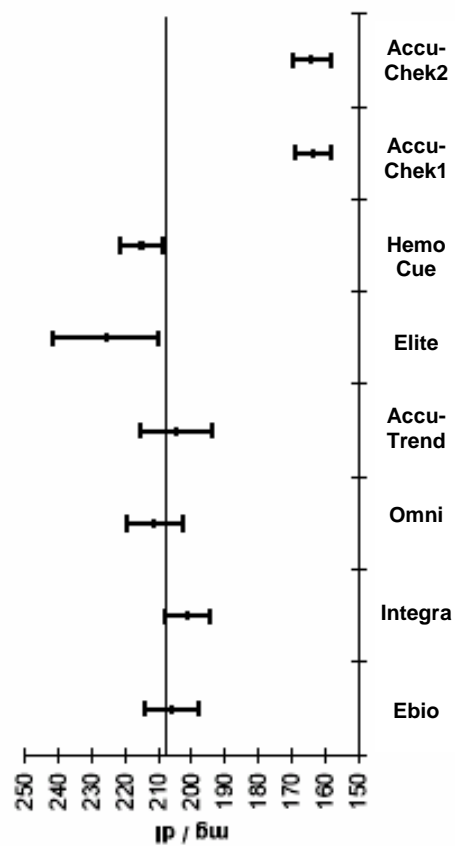
SW	Sollwert [mg/dl]
MW	Mittelwert [mg/dl] (arithmetisch gemittelte Messergebnisse, berechnet aus den täglichen Mittelwerten der Doppelbestimmungen)
SD	Standardabweichung [mg/dl]
Min	Minimum [mg/dl]
Max	Maximum [mg/dl]
VK%	Variationskoeffizient (Maß für die Präzision von-Tag-zu-Tag entsprechend der relativen Streuung um den MW)
△%	prozentuale Abweichung des MW vom SW
N	Anzahl der Glukosemessungen

Abbildungen 10 a-d: Mittelwerte und Standardabweichungen der gemessenen Glukosekonzentrationen der Standard- und Kontroll-Lösungen
 (Sollwert ist als durchgezogene Linie eingezeichnet)

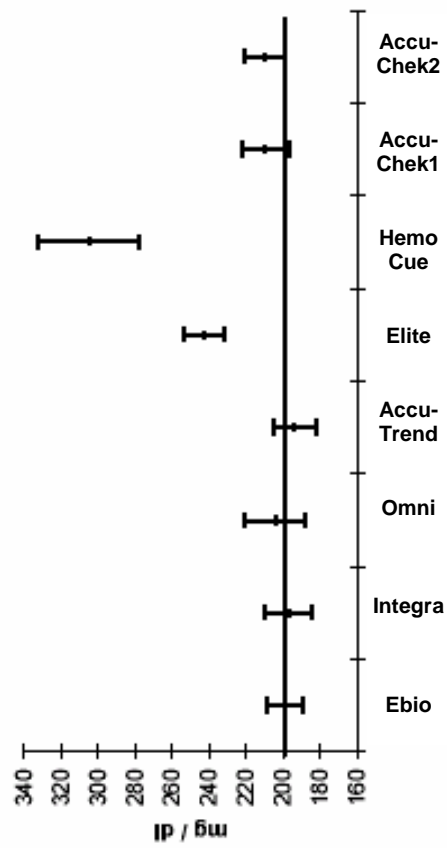
10a: Glu-NIST (Sollwert 200 ma/dl)



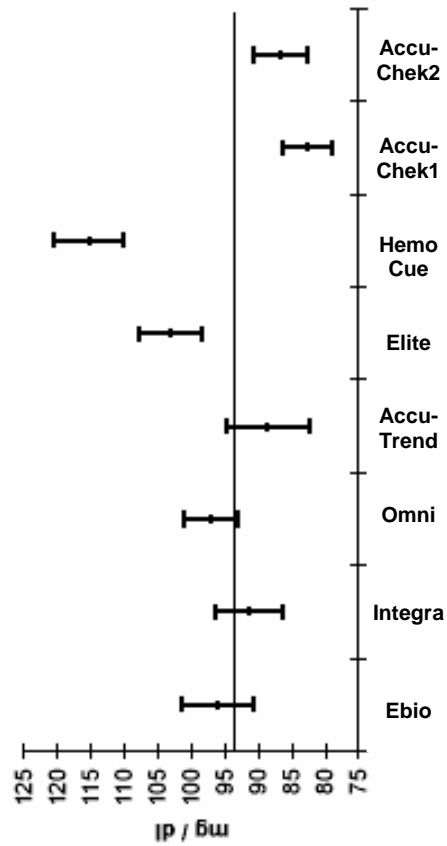
10c: Val A (Sollwert 208 ma/dl)



10b: Glu-Alb (Sollwert 200 ma/dl)



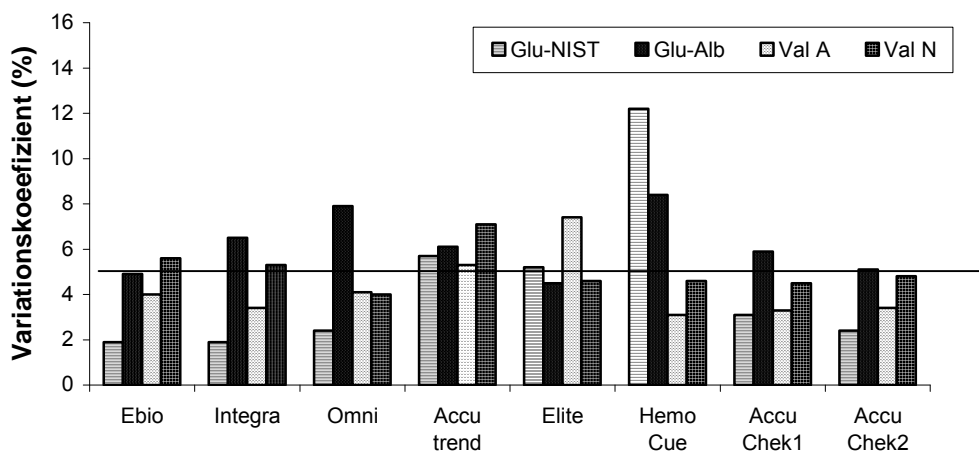
10d: Val N (Sollwert 93.9 ma/dl)



In den Abbildungen 10 a - d wird deutlich erkennbar, dass die Messungen von Ebio, Integra, Omni und Accutrend im Akzeptanzbereich lagen (Mittelwerte lagen um maximal 5% vom Sollwert entfernt). Elite maß die Standard- und Kontroll-Lösungen im Durchschnitt zu hoch, HemoCue bis auf Val A ebenfalls. Die beiden Accu-Cheks maßen bis auf den Glu-Alb-Standard im Durchschnitt zu niedrig.

Die Variationskoeffizienten werden in Abbildung 11 als Säulendiagramme dargestellt mit Einzeichnung einer 5%-Linie.

Abbildung 11: Präzision von-Tag- zu-Tag



Variationskoeffizienten, die unter 5% lagen (Tabelle 12, Abbildung 11), zeigten sich bei Ebio (Glu-NIST, Val A), Integra (Glu-NIST, Val A), Omni (Glu-NIST, Val A, Val N), Elite (Glu-Alb, Val N), HemoCue (Val A, Val N), Accu-Chek 1 (Glu-NIST, Val A, Val N) und Accu-Chek 2 (Glu-NIST, Val A, Val N). Anhand der Abbildung 11 fiel der Variationskoeffizient des HemoCue beim Glu-NIST mit 12,2% aus der Reihe, was dahingehend zu interpretieren ist, dass HemoCue für Glukosemessungen in proteinfreien Lösungen nicht geeignet ist. Zusammenfassend zeigte sich eine Impräzision (angegeben durch VK%) bei den Kontroll-Lösungen (Val A, Val N) zwischen 3,1 und 7,4%, bei den Standardlösungen (Glu-NIST, Glu-Alb) zwischen 1,9 und 12,2%, bzw. zwischen 1,9 und 8,4%, wenn man HemoCue Glu-Alb außer Betracht lässt.

4.1.2 Tages-bezogene Korrekturfaktoren

Die Berechnung der Tages-bezogenen Korrekturfaktoren wurde in Kapitel 3.2.1.1 beschrieben. In dieser Auswertung wurden die Glu-NIST-bezogenen Tages-Korrekturfaktoren berechnet und angewandt. Tabelle 13 enthält die über den Studienzeitraum gemittelten Tages-bezogenen Korrekturfaktoren und deren Minimum und Maximum.

Tabelle 13: Tages-bezogene Korrekturfaktoren, ermittelt anhand des Glu-NIST

Glukometer	Tages-bezogene Korrekturfaktoren		
	Durchschnitt	Minimum	Maximum
Ebio	0,98	0,90	1,04
Integra	1,03	0,93	1,23
Omni	1,01	0,95	1,13
Accutrend	0,99	0,87	1,18
HemoCue	0,88	0,72	1,18
Elite	0,85	0,73	1,17
Accucheck1	1,17	0,92	1,82
Accucheck2	1,15	1,04	1,41

Aus der Tabelle 13 ist erkennbar, dass die durchschnittlich ermittelten Tages-bezogenen Korrekturfaktoren von Ebio, Integra, Omni und Accutrend dicht bei 1,0 lagen, und um maximal 3% abwichen. Die unkorrigierten Messwerte dieser Glukometer blieben infolgedessen nach Tages-Korrektur quasi unverändert. Die anderen Geräte wiesen Tages-bezogene Korrekturfaktoren auf, die um mehr als 10% von 1,0 abwichen. Die unkorrigierten Messwerte des HemoCue und Elite wurden durch deren Tages-bezogenen Korrekturfaktoren von 0,88 und 0,85 respektive nach unten korrigiert, wohingegen die der Accu-Cheks durch deren Korrekturfaktoren von 1,17 und 1,15 nach oben korrigiert wurden.

Da der Tages-bezogene Korrekturfaktor des Ebio bei 0,98 lag, was unveränderte Werte nach Korrektur zur Folge hatte, wurde beschlossen, bei den intersystemischen Vergleichen (Testgerät-Messwerte vs. Ebio-Messwerte) ausschließlich die Tages-korrigierten Ebio-Messwerte als Referenzwerte zu nehmen.

4.2 MESSUNGEN DES MENSCHLICHEN PROBENMATERIALS

4.2.1 Rate an falsch positiven und falsch negativen Messwerten

Die Rate an falsch positiven und falsch negativen Messergebnissen (siehe Kapitel 3.2.2.3) lag bei den intrasystemischen Vergleichen (Tabelle 14) bei fast allen Testgeräten unter 5%. In wenigen Fällen lag sie minimal darüber, die maximale Rate lag bei 6,7% (venöses Nüchternblut von Elite). Im 2h-Postprandialblut war die Rate an falsch positiven und falsch negativen Messergebnissen niedriger als im Nüchternblut.

Tabelle 14: Anzahl an falsch positiven und falsch negativen Messungen bei den „intrasystemischen“ Vergleichen bezüglich der Entscheidungsgrenze „diabetisch / nicht-diabetisch“

Glukometer	PS	Zeitpunkt Null			Zeitpunkt 120 Min		
		n	Inkongruenz um EG absolut [prozentual]		n	Inkongruenz um EG absolut [prozentual]	
Ebio (Pop-Kollektiv)	CB	(87)	3	[3,4]	(89)	4	[4,5]
Ebio (Pop-Kollektiv)	VB	(111)	3	[2,7]	(109)	2	[1,8]
Ebio (Pop-Kollektiv)	VP	(114)	4	[3,5]	(108)	1	[0,9]
Ebio (Stud-Kollektiv)	CB	(69)	2	[2,9]	(73)	3	[4,1]
Ebio (Stud-Kollektiv)	VB	(76)	2	[2,6]	(75)	1	[1,3]
Ebio (Stud-Kollektiv)	VP	(77)	3	[3,9]	(75)	1	[1,3]
Integra	CB	(42)	1	[2,4]	(42)	1	[2,4]
Integra	VB	(45)	2	[4,4]	(44)	0	.
Integra	VP	(46)	0	.	(46)	1	[2,2]
Omni	CB	(36)	0	.	(40)	0	.
Omni	VB	(69)	4	[5,8]	(66)	2	[3]
Omni	VP	(64)	0	.	(61)	2	[3,3]
Accutrend	CB	(34)	2	[5,9]	(43)	0	.
Accutrend	VB	(51)	0	.	(52)	1	[1,9]
Accutrend	VP	(53)	1	[1,9]	(51)	0	.
HemoCue	CB	(27)	1	[3,7]	(29)	0	.
HemoCue	VB	(42)	2	[4,8]	(42)	1	[2,4]
HemoCue	VP	(42)	2	[4,8]	(41)	0	.
Elite	CB	(32)	2	[6,3]	(36)	1	[2,8]
Elite	VB	(45)	3	[6,7]	(46)	0	.
Elite	VP	(47)	3	[6,4]	(45)	1	[2,2]
AccucheK1	CB	(27)	1	[3,7]	(33)	1	[3]
AccucheK1	VB	(40)	0	.	(39)	2	[5,1]
AccucheK1	VP	(40)	0	.	(38)	1	[2,6]
AccucheK2	CB	(32)	0	.	(38)	0	.
AccucheK2	VB	(44)	2	[4,5]	(43)	0	.
AccucheK2	VP	(44)	1	[2,3]	(42)	0	.

- EG im/md = Entscheidungsgrenze zwischen glukoseintolerant (intermediär = im) und diabetisch (mit Diabetes = md)
- PS = Probensystem (CB = kapilläres Blut, VB = venöses Blut, VP = venöses Plasma)
- n = Anzahl der Wertepaare (entsprechend der Anzahl an Doppelmessungen)
- Inkongruenz absolut = Summe der falsch-positiven und falsch-negativen Konzentrationsmessungen
- Inkongruenz prozentual = Rate an Inkongruenzen in Bezug auf die Anzahl der Wertepaare
- Pop-Kollektiv = Populationskollektiv (Zeitraum 04/94 bis 11/01, n = 168, Ebio)
- Stud-Kollektiv = Studienkollektiv (Zeitraum 06/99 bis 09/01, n = 78, Testgerät vs. Ebio)

Bei den intersystemischen Vergleichen lag, wie in Tabelle 15 zu erkennen ist, die Rate an Falschmessungen bei den meisten Glukometer über 5%.

Tabelle 15: Anzahl an falsch positiven und falsch negativen Messungen bei den „intersystemischen“ Vergleichen bezüglich der Entscheidungsgrenze „diabetisch / nicht-diabetisch“

	Nüchternzeitpunkt				Postprandialzeitpunkt				
	Glukometer	(n)	fn	fp	f(n+p)	(n)	fn	fp	f(n+p)
Kapillarblut	Integra	(45)	0 .	0 .	0 .	(45)	0 .	2 [4,4]	2 [4,4]
	Integra pTk	(45)	2 [4,4]	1 [2,2]	3 [6,7]	(45)	0 .	2 [4,4]	2 [4,4]
	Omni	(62)	0 .	26 [41,9]	26 [41,9]	(67)	0 .	12 [17,9]	12 [17,9]
	Omni pTk	(62)	0 .	27 [43,5]	27 [43,5]	(67)	0 .	11 [16,4]	11 [16,4]
	Elite	(48)	0 .	6 [12,5]	6 [12,5]	(47)	0 .	7 [14,9]	7 [14,9]
	Elite pTk	(48)	1 [2,1]	0 .	1 [2,1]	(47)	0 .	3 [6,4]	3 [6,4]
	HemoCue	(44)	0 .	1 [2,3]	1 [2,3]	(43)	0 .	2 [4,7]	2 [4,7]
	HemoCue pTk	(44)	4 [9,1]	0 .	4 [9,1]	(43)	0 .	2 [4,7]	2 [4,7]
	Accutrend	(52)	0 .	3 [5,8]	3 [5,8]	(54)	0 .	6 [11,1]	6 [11,1]
	Accutrend pTk	(52)	0 .	6 [11,5]	6 [11,5]	(54)	0 .	5 [9,3]	5 [9,3]
	Accu-chek1	(40)	0 .	2 [5]	2 [5]	(40)	0 .	3 [7,5]	3 [7,5]
	Accu-chek1 pTk	(40)	0 .	15 [37,5]	15 [37,5]	(40)	0 .	8 [20]	8 [20]
	Accu-chek2	(45)	0 .	2 [4,4]	2 [4,4]	(45)	0 .	4 [8,9]	4 [8,9]
	Accu-chek2 pTk	(45)	0 .	19 [42,2]	19 [42,2]	(45)	0 .	7 [15,6]	7 [15,6]
	Venöses Vollblut	Glukometer	(n)	fn	fp	f(n+p)	(n)	fn	fp
Integra		(46)	0 .	0 .	0 .	(44)	1 [2,2]	1 [2,2]	2 [4,3]
Integra pTk		(46)	1 [2,3]	1 [2,3]	2 [4,5]	(44)	1 [2,3]	2 [4,5]	3 [6,8]
Omni		(71)	0 .	18 [25,4]	18 [25,4]	(68)	0 .	20 [28,2]	20 [28,2]
Omni pTk		(71)	0 .	7 [10,3]	7 [10,3]	(68)	0 .	9 [13,2]	9 [13,2]
Elite		(48)	2 [4,4]	1 [2,2]	3 [6,7]	(46)	5 [11,1]	0 .	5 [11,1]
Elite pTk		(48)	0 .	1 [2,3]	1 [2,3]	(46)	2 [4,7]	2 [4,7]	4 [9,3]
HemoCue		(45)	0 .	2 [3,8]	2 [3,8]	(43)	1 [1,9]	3 [5,7]	4 [7,5]
HemoCue pTk		(45)	0 .	1 [1,9]	1 [1,9]	(43)	0 .	2 [3,8]	2 [3,8]
Accutrend		(53)	0 .	3 [7,3]	3 [7,3]	(52)	0 .	14 [34,1]	14 [34,1]
Accutrend pTk		(53)	0 .	1 [2,6]	1 [2,6]	(52)	0 .	7 [17,9]	7 [17,9]
Accu-chek1		(41)	0 .	2 [4,4]	2 [4,4]	(39)	0 .	19 [42,2]	19 [42,2]
Accu-chek1 pTk		(41)	0 .	4 [9,3]	4 [9,3]	(39)	0 .	5 [11,6]	5 [11,6]
Accu-chek2		(45)	0 .	7 [14,6]	7 [14,6]	(43)	2 [4,2]	0 .	2 [4,2]
Accu-chek2 pTk		(45)	0 .	4 [8,7]	4 [8,7]	(43)	2 [4,3]	2 [4,3]	4 [8,7]
Venöses Plasma	Glukometer	(n)	fn	fp	f(n+p)	(n)	fn	fp	f(n+p)
	Integra	(47)	1 [2,1]	0 .	1 [2,1]	(46)	1 [2,1]	0 .	1 [2,1]
	Integra pTk	(47)	1 [2,2]	1 [2,2]	2 [4,3]	(46)	1 [2,2]	0 .	1 [2,2]
	Omni	(70)	1 [1,4]	1 [1,4]	2 [2,9]	(68)	1 [1,4]	1 [1,4]	2 [2,9]
	Omni pTk	(70)	0 .	1 [1,5]	1 [1,5]	(68)	0 .	2 [2,9]	2 [2,9]
	Elite	(48)	0 .	2 [4,4]	2 [4,4]	(46)	4 [8,9]	2 [4,4]	6 [13,3]
	Elite pTk	(48)	0 .	1 [2,3]	1 [2,3]	(46)	2 [4,7]	1 [2,3]	3 [7]
	HemoCue	(45)	0 .	1 [1,9]	1 [1,9]	(43)	1 [1,9]	2 [3,7]	3 [5,6]
	HemoCue pTk	(45)	0 .	0 .	0 .	(43)	0 .	1 [1,9]	1 [1,9]
	Accutrend	(54)	2 [4,9]	1 [2,4]	3 [7,3]	(52)	0 .	6 [14,6]	6 [14,6]
	Accutrend pTk	(54)	5 [12,8]	0 .	5 [12,8]	(52)	0 .	2 [5,1]	2 [5,1]
	Accu-chek1	(41)	3 [6,7]	0 .	3 [6,7]	(39)	0 .	7 [15,6]	7 [15,6]
	Accu-chek1 pTk	(41)	4 [9,3]	0 .	4 [9,3]	(39)	0 .	2 [4,7]	2 [4,7]
	Accu-chek2	(45)	0 .	19 [39,6]	19 [39,6]	(43)	0 .	1 [2,1]	1 [2,1]
	Accu-chek2 pTk	(45)	0 .	7 [15,2]	7 [15,2]	(43)	1 [2,2]	0 .	1 [2,2]

n = Anzahl der Wertepaare

fn = falsch negative Messergebnisse [in Prozent an der Anzahl der Wertepaare]

fp = falsch positive Messergebnisse [in Prozent an der Anzahl der Wertepaare]

f(n+p) = falsch negative und falsch positive Messergebnisse [in Prozent an der Anzahl der Wertepaare]

pTk = pro-Tag-korrigiert (auf den Glu-NIST-Standard bezogene Messergebnisse)

4.2.2 Konventioneller Methodenvergleich

Für alle intrasystemischen Vergleichen wurde im Rahmen der Passing/Bablok-Regressionsanalysen mittels Cusum-Test eine lineare Beziehung nachgewiesen. Die Ausgleichsgeraden lagen visuell auf der Winkelhalbierenden. Die Steigungen entsprachen demzufolge 1,0 oder wichen kaum von 1,0 ab, und die Achsen wurden bei 0 oder nur minimal abweichend von 0 geschnitten. Auf tabellarische und graphische Darstellung wurde hier verzichtet.

Auch bei den intersystemischen Vergleichen wies der Cusum-Test keine signifikante Abweichung von der Linearität auf. Allerdings lagen bei visueller Inspektion die Ausgleichsgeraden nicht auf der Winkelhalbierenden, sondern waren entweder parallel verschoben ($\alpha \neq 0, \beta = 1$), wiesen eine Steigungsabweichung auf ($\alpha = 0, \beta \neq 1$), oder – in den meisten Fällen – waren sowohl verschoben als auch steigungsabweichend ($\alpha \neq 0, \beta \neq 1$). In den Tabellen 16 und 17 sind die Ergebnisse der Regressionsanalysen aller unkorrigierten und Tages-korrigierten intersystemischen Vergleiche erfasst. Die Tabellen enthalten die Angaben zu Anzahl n der Wertepaare, Achsenabschnitt β , Steigung α , Korrelationskoeffizient r , md_{95} , die mittlere Konzentration der X-Werte und der gepaarten Y-Werte und die prozentuale Abweichung der mittleren Konzentration Y von der mittleren Konzentration X („MW-Abweichung“).

Tabelle 16: Intersystemische Vergleiche zum Nüchternzeitpunkt. Ergebnisse der linearen Regressionsanalysen nach Passing-Bablok (Testwert-Messungen Y vs. Tages-korrigierte Ebio-Messungen X)

Glukometer	(n)	β	α	r	md(95)	x-quer	y-quer	MW-Abw
Kapillarblut								
Integra	(45)	1,019	- 5,36	0,977	5,504	96,0	92,7	- 3,4*
Integra, pTk	(45)	1,034	- 6,62	0,926	9,330	96,0	94,6	- 1,4
Omni	(62)	1,497*	- 25,16*	0,915	8,273	93,2	116,5	23,1*
Omni, pTk	(62)	1,507*	- 24,36*	0,902	8,893	93,5	116,5	24,6*
Elite	(48)	1,330*	- 20,84*	0,957	7,853	96,7	105,9	9,5*
Elite, pTk	(48)	1,167*	- 21,84*	0,927	8,780	96,7	92,2	- 4,6*
HemoCue	(44)	1,138	- 12,69	0,942	7,869	96,1	97,0	0,9
HemoCue, pTk	(44)	1,149	- 27,56*	0,797	15,045	96,1	85,7	-10,8*
Accutrend	(52)	1,032	+ 2,57	0,913	12,435	95,1	102,3	7,6*
Accutrend, pTk	(52)	1,059	+ 3,09	0,919	9,358	95,1	103,5	8,7*
Accu-Chek 1	(40)	1,125*	- 6,05	0,966	7,888	96,8	102,5	5,9*
Accu-Chek 1, pTk	(40)	1,382*	- 15,18*	0,962	9,824	96,8	117,9	21,8*
Accu-Chek 2	(45)	1,100	+ 1,72	0,969	6,848	96,3	103,5	7,4*
Accu-Chek 2, pTk	(45)	1,222	+ 0,48	0,967	7,138	96,3	118,1	22,6*
Venöses Vollblut								
Integra	(46)	1,032	- 7,17	0,974	5,850	92,4	88,7	- 4,0*
Integra, pTk	(46)	1,008	- 5,99	0,934	9,622	92,4	90,3	- 2,3*
Omni	(71)	1,349	- 12,61	0,920	10,912	94,2	113,1	20,1*
Omni, pTk	(71)	1,364*	- 14,31	0,911	10,706	94,2	114,4	21,5*
Elite	(48)	1,240*	- 10,84	0,915	9,545	95,4	107,2	12,3*
Elite, pTk	(48)	1,168*	- 17,15	0,902	11,973	95,4	93,3	- 2,2
HemoCue	(45)	1,043	- 6,80	0,910	9,858	95,4	92,5	- 3,1*
HemoCue, pTk	(45)	0,975	- 12,51	0,813	13,897	95,4	81,8	-14,3*
Accutrend	(53)	1,008	- 0,14	0,921	11,319	94,2	96,6	2,6*
Accutrend, pTk	(53)	1,088	- 3,68	0,927	9,558	94,2	97,7	3,7*
Accu-Chek 1	(41)	1,052	- 0,32	0,966	6,194	94,6	99,4	5,1*
Accu-Chek 1, pTk	(41)	1,231*	- 2,66	0,954	10,623	94,6	114,3	20,8*
Accu-Chek 2	(45)	1,084	- 0,84	0,952	8,573	95,3	101,9	6,9*
Accu-Chek 2, pTk	(45)	1,226*	- 0,08	0,950	8,439	95,3	116,3	22,1*
Venöses Plasma								
Integra	(47)	1,061	- 8,63	0,945	11,025	106,4	102,7	- 3,5*
Integra, pTk	(47)	0,973	- 0,08	0,985	5,617	106,4	104,0	- 2,3*
Omni	(70)	1,030	- 4,34	0,955	8,703	107,9	106,9	- 0,9
Omni, pTk	(70)	1,018	- 2,12	0,947	9,139	107,9	108,1	0,2
Elite	(48)	1,380*	- 19,89*	0,963	9,550	110,3	132,4	20,1*
Elite, pTk	(48)	1,244*	- 22,90*	0,962	7,526	110,3	115,1	4,4*
HemoCue	(45)	0,950	+13,49	0,923	8,791	109,7	117,5	7,0*
HemoCue, pTk	(45)	1,048	- 14,43	0,758	19,131	109,7	104,0	- 5,2*
Accutrend	(54)	1,075	- 12,03	0,926	13,530	108,4	106,3	- 2,0*
Accutrend, pTk	(54)	1,038	- 6,37	0,933	10,349	108,4	107,1	- 1,2
Accu-Chek 1	(41)	0,881*	+ 7,10	0,966	5,599	109,9	103,4	- 6,0*
Accu-Chek 1, pTk	(41)	1,073	+ 0,40	0,970	6,805	109,9	118,8	8,0*
Accu-Chek 2	(45)	0,979	+ 1,61	0,950	7,762	109,9	107,7	- 2,0*
Accu-Chek 2, pTk	(45)	1,060	+ 5,84	0,953	9,391	109,9	122,8	11,8*

(n) = Anzahl der Wertepaare

β = Steigung

α = Achsenabschnitt

r = Korrelationskoeffizient

md(95) = mean 95%-deviation = Maß der Streuung um die Ausgleichsgerade

x-quer = mittlere Konzentrationen der Referenzwerte (x) in mg/dl

y-quer = mittlere Konzentrationen der Testwerte (y) in mg/dl

MW-Abw = Mittelwert-Abweichung = prozentuale Abweichung des y-quer von x-quer in mg/dl (berechnet: $f(x) = (y - x) * 100 / x$)

pTk = pro-Tag-korrigiert (Bezug der rohen Messungen auf den wässrigen Glukosestandard Glu-NIST)

* = signifikante Abweichung, mittels zweiseitigem t-Test errechnet ($p < 0,05$)

Tabelle 17: Intersystemische Vergleiche zum 2h-Postprandial-Zeitpunkt. Ergebnisse der linearen Regressionsanalysen nach Passing-Bablok (Testwert-Messungen Y versus Tages-korrigierte Ebio-Messungen X)

Glukometer	(n)	β	α	r	md(95)	x-quer	y-quer	MW-Abw
Kapillarblut								
Integra	(45)	1,102	- 16,57	0,986	8,561	140,0	136,7	- 2,4*
Integra, pTk	(45)	1,035	- 7,83	0,968	11,750	140,0	139,5	- 0,4
Omni	(67)	1,360	- 17,63*	0,964	18,130	142,2	176,1	23,9*
Omni, pTk	(67)	1,364*	- 17,68*	0,954	18,778	142,2	178,3	25,4*
Elite	(47)	1,289	- 23,35*	0,970	14,704	141,8	159,5	12,5*
Elite, pTk	(47)	1,228	- 34,52*	0,958	15,501	141,8	139,6	- 1,6
HemoCue	(43)	1,168	- 19,37*	0,971	16,895	141,1	147,7	4,6*
HemoCue, pTk	(43)	1,196	- 45,03*	0,919	26,958	141,1	131,8	- 6,6*
Accutrend	(54)	1,119	- 9,55	0,968	13,953	140,1	148,8	6,2*
Accutrend, pTk	(54)	1,163	- 14,48	0,972	12,472	140,1	149,7	6,8*
Accu-Chek 1	(40)	1,145	- 10,71	0,979	10,491	142,3	152,2	7,0*
Accu-Chek 1, pTk	(40)	1,426	- 27,45*	0,975	14,534	142,3	175,3	23,2*
Accu-Chek 2	(45)	1,165	- 11,94*	0,976	13,328	143,0	153,7	7,5*
Accu-Chek 2, pTk	(45)	1,258	- 4,50	0,962	15,537	143,0	174,8	22,2*
Venöses Vollblut								
Integra	(44)	1,024	- 6,02	0,991	6,934	124,9	121,5	- 2,7*
Integra, pTk	(44)	1,005	- 3,49	0,979	10,137	124,9	123,8	- 0,9
Omni	(68)	1,271	- 4,88	0,977	13,932	131,2	158,7	20,9*
Omni, pTk	(68)	1,280	- 4,48	0,974	15,805	131,2	160,6	22,4*
Elite	(46)	1,174	- 1,89	0,983	11,926	131,0	150,9	15,1*
Elite, pTk	(46)	1,056	- 4,38	0,959	18,820	131,0	131,6	0,5
HemoCue	(43)	1,153	- 18,96*	0,988	9,264	129,0	129,2	0,1
HemoCue, pTk	(43)	1,072	- 26,45*	0,952	21,355	129,0	115,3	-10,7*
Accutrend	(52)	1,051	- 4,70	0,985	9,673	128,1	129,1	0,7
Accutrend, pTk	(52)	1,037	- 2,97	0,986	10,908	128,1	130,0	1,4
Accu-Chek 1	(39)	1,057	- 2,72	0,988	12,422	129,2	131,8	2,0*
Accu-Chek 1, pTk	(39)	1,290	- 11,77*	0,982	14,635	129,2	152,2	17,8*
Accu-Chek 2	(43)	1,023	+ 1,96	0,980	12,558	131,4	136,0	3,5*
Accu-Chek 2, pTk	(43)	1,141	+ 4,82	0,977	11,848	131,4	154,9	17,9*
Venöses Plasma								
Integra	(46)	1,047	- 8,30	0,971	17,485	145,1	139,7	- 3,7*
Integra, pTk	(46)	0,988	- 2,12	0,994	5,567	145,1	139,7	- 2,6*
Omni	(68)	1,050	- 7,79*	0,988	9,335	151,2	151,0	- 0,1
Omni, pTk	(68)	1,058	- 6,86*	0,985	12,459	151,2	152,8	1,1
Elite	(46)	1,275	- 13,18	0,974	18,780	152,3	180,0	18,2*
Elite, pTk	(46)	1,077	- 8,31	0,967	14,191	152,3	156,7	2,9
HemoCue	(43)	0,957	+12,81*	0,983	14,519	149,7	156,0	4,2*
HemoCue, pTk	(43)	0,914	- 0,29	0,935	21,903	149,7	138,9	- 7,2*
Accutrend	(52)	1,084	- 16,60*	0,987	10,828	149,1	144,5	- 3,1*
Accutrend, pTk	(52)	1,080	- 15,59*	0,989	11,154	149,1	145,4	- 2,5*
Accu-Chek 1	(39)	0,883	+ 3,15	0,984	12,718	149,3	135,2	- 9,5*
Accu-Chek 1, pTk	(39)	1,080	- 4,52	0,989	9,357	149,3	155,9	4,4*
Accu-Chek 2	(43)	0,847	+12,12	0,971	16,298	152,5	139,8	- 8,3*
Accu-Chek 2, pTk	(43)	0,961	+14,27	0,973	15,806	152,5	159,1	4,3*

(n) = Anzahl der Wertepaare

β = Steigung

α = Achsenabschnitt

r = Korrelationskoeffizient

md(95) = mean 95%-deviation = Maß der Streuung um die Ausgleichsgerade

x-quer = mittlere Konzentrationen der Referenzwerte (x) in mg/dl

y-quer = mittlere Konzentrationen der Testwerte (y) in mg/dl

MW-Abw = Mittelwert-Abweichung = prozentuale Abweichung des y-quer von x-quer in mg/dl (berechnet: $f(x) = (y - x) * 100 / x$)

pTk = pro-Tag-korrigiert (Bezug der rohen Messungen auf den wässrigen Glukosestandard Glu-NIST)

* = signifikante Abweichung, mittels zweiseitigem t-Test errechnet ($p < 0,05$)

Die Graphiken der Regressionsanalysen mit den dazugehörigen Differenzenplots sind in den Abbildungen 12 – 25 im Anhang hinterlegt. Anhand der MW-Abweichungen wurde überprüft, in welchem Rahmen die Testgeräte mit ihren Messungen von denen des Ebio-Gerätes abwichen, inwieweit sie also innerhalb oder außerhalb eines Akzeptanzbereiches von 5% lagen.

Die Messungen des Integras lagen alle im Akzeptanzbereich, sie wiesen maximale MW-Abweichungen von 5% im Vergleich zu den Ebio-Messungen auf. Die Vollblut-Messungen des Omni lagen weit außerhalb des Akzeptanzbereiches (MW-Abweichungen zwischen 20,1% bis 23,9%), die VP-Messungen hingegen lagen im Akzeptanzbereich. Alle Messungen des Elite lagen außerhalb des Akzeptanzbereiches (MW-Abweichungen zwischen 9,5% bis 20,1%). Die Vollblut- und 2h-postprandialen VP-Messungen des HemoCue lagen im Akzeptanzbereich, die Nüchtern-VP-Messungen wiesen eine geringe MW-Abweichung von 7,0% auf. Die CB-Messungen des Accutrends lagen etwas außerhalb des Akzeptanzbereiches (MW-Abweichungen zwischen 6,2% und 7,6%), die VB- und VP-Messungen hingegen im Akzeptanzbereich. Die CB- und Nüchtern-VB-Messungen der beiden Accu-Cheks und die Nüchtern-VP-Messungen des Accu-Chek2 lagen etwas außerhalb des Akzeptanzbereiches (MW-Abweichungen zwischen 5,1% und 7,5%), die 2h-postprandialen VB-Messungen der beiden Accu-Cheks und die 2h-postprandialen VP-Messungen des Acc-Chek2 lagen im Akzeptanzbereich.

Entsprechend der Tages-bezogenen Korrekturfaktoren (Tabelle 13, Kapitel 4.1.1) wurden die Messergebnisse des Integra, Omni und Accutrend quasi nicht verändert, die des HemoCue und Elite nach unten und die der Accu-Cheks nach oben korrigiert. Infolgedessen blieben die Messergebnisse des Integra im Akzeptanzbereich (MW-Abweichungen unter 5%), die Vollblut-Messungen des Omni weit außerhalb des Akzeptanzbereiches (MW-Abweichungen zwischen 21,5% bis 25,4%), die VP-Messungen des Omni im Akzeptanzbereich, die CB-Messungen des Accutrend etwas außerhalb des Akzeptanzbereiches (MW-Abweichungen zwischen 6,8% bis 8,7%), die VB- und VP-Messungen des Accutrend im Akzeptanzbereich. Die Messungen des HemoCue lagen nach Tages-Korrektur außerhalb des Akzeptanzbereiches (MW-Abweichungen zwischen -5,2% und -14,3%), die Messungen des Elite lagen nach Tages-Korrektur im Akzeptanzbereich. Die CB- und VB-Messungen der Accu-Cheks lagen nach Tages-Korrektur weit außerhalb des Akzeptanzbereiches (MW-Abweichungen zwischen 17,8% bis 23,2%), die VP-Messungen zum Nüchtern-Entnahmezeitpunkt etwas außerhalb (MW-Abweichungen zwischen 8,0% bis 11,8%) und die VP-Messungen zum 2h-postprandialen Entnahmezeitpunkt innerhalb des Akzeptanzbereiches.

Die md_{95} -Werte (Maß für die Streuung um die Ausgleichsgerade) waren zum 2h-postprandial-Zeitpunkt deutlich höher als zum Nüchtern-Zeitpunkt: nüchtern lagen die md_{95} -Werte durchschnittlich um 8,8 mg/dl (Minimum = 5,5, Maximum = 13,53), postprandial lagen die md_{95} -Werte durchschnittlich um 13,0 mg/dl (Minimum = 6,9, Maximum = 18,8). Es war kein Trend zu erkennen, ob die md_{95} -Werte einer speziellen Probenart höher waren als die einer anderen (siehe Tabellen 16 und 17).

4.2.3 Diskordanzraten-Test

4.2.3.1 Diskordanzraten der intrasystemischen Vergleiche

Die Tabellen 18 und 19 enthalten die Ergebnisse des DR-Testes der intrasystemischen Vergleiche, also die intrasystemischen Diskordanzraten. Darin sind die Gesamt-DR, die Anzahl der Wertepaare und die Größe der jeweiligen Populationskollektive erfasst.

Tabelle 18: Gesamt-Diskordanzraten (%) der intrasystemischen Vergleiche zum Nüchternzeitpunkt (Erklärungen der Abkürzungen, siehe Tabelle 18)

	PS	EG	Ebio*	INT	O	AT	HC	E	AC1	AC2	Pop-Koll**
Anzahl (n) der Wertepaare	CB ₂ /CB ₁ ***	110	85	41	36	34	27	32	26	32	150
	VB ₂ /VB ₁	110	109	45	68	50	41	44	39	43	167
	VP ₂ /VP ₁	126	111	46	63	52	41	46	39	43	160
Gesamt-DR	CB ₂ /CB ₁	110	3,0	3,6	3,5	4,0	4,2	5,9	2,2	2,9	
	VB ₂ /VB ₁	110	2,5	3,4	2,8	2,4	3,4	3,6	2,3	2,1	
	VP ₂ /VP ₁	126	1,5	1,8	1,8	2,0	2,2	3,1	2,1	2,1	

Tabelle 19: Gesamt-Diskordanzraten (%) der intrasystemischen Vergleiche zum 2h-Postprandial-Zeitpunkt

	PS	EG	Ebio*	INT	O	AT	HC	E	AC1	AC2	Pop-Koll**
Anzahl (n) der Wertepaare	CB ₂ /CB ₁	200	86	41	39	42	29	35	32	37	168
	VB ₂ /VB ₁	180	106	43	65	51	41	45	38	42	146
	VP ₂ /VP ₁	200	105	45	60	50	40	44	37	41	156
Gesamt-DR	CB ₂ /CB ₁	200	1,7	2,4	1,5	1,8	1,4	2,9	2,0	1,3	
	VB ₂ /VB ₁	180	1,6	1,2	1,9	1,3	1,5	1,8	1,5	1,1	
	VP ₂ /VP ₁	200	1,3	1,0	1,2	0,9	1,0	2,4	1,4	0,8	

PS = Probensystem, EG = Entscheidungsgrenze, ZP = Zeitpunkt,

CB = Kapillarblut, VB = Venöses Blut, VP = Venöses Plasma

INT = Integra, O = Omni, AT = Akkutrend, HC = HemoCue, E = Elite, AC1 = Accu-Chek1, AC2 = Accu-Chek2

*) analytische Wertepaare der Ebio-Messungen des Populationskollektiv

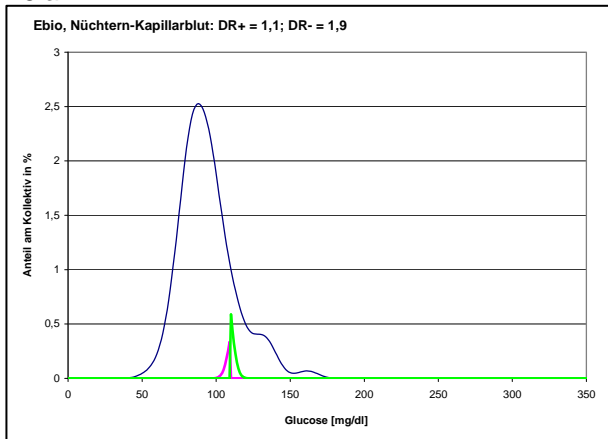
**) Anzahl der Messungen des jeweiligen PS des Populationskollektiv

***) Zweitmessung versus Erstmessung des jeweiligen PS

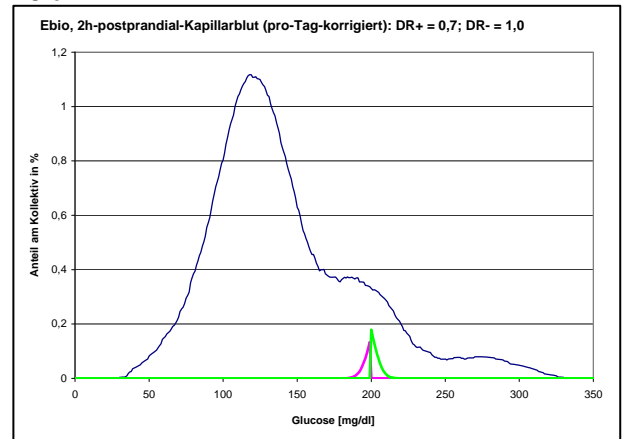
In den Abbildungen 26 a - f sind die DR-Graphiken des intrasystemischen Vergleiches von Ebio abgebildet. Auf die graphische Darstellung der restlichen intrasystemischen Vergleiche wurde verzichtet.

Abbildungen 26 a – f: Intrasystemische Vergleiche von Ebio. Graphische Darstellung der Ergebnisse des Diskordanzratentests

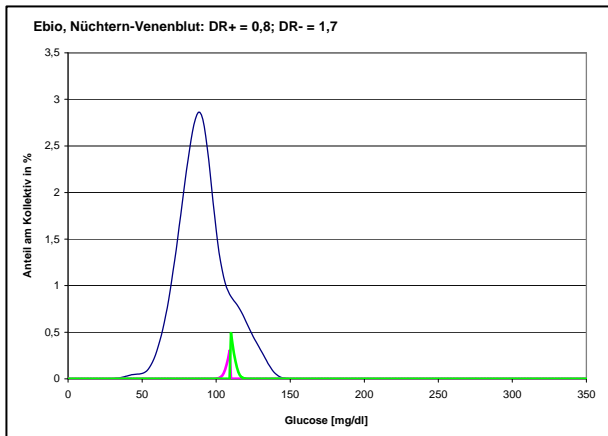
26 a:



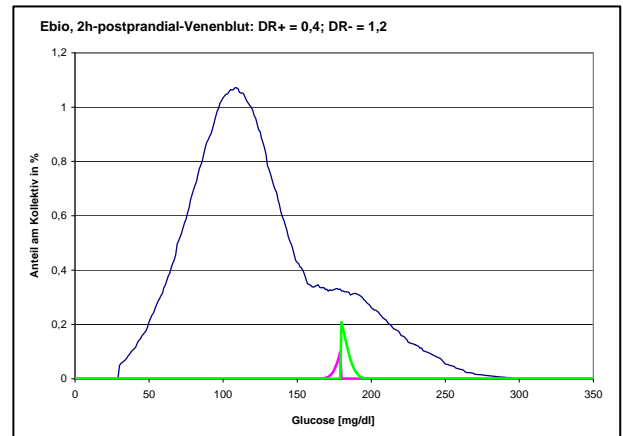
26 d:



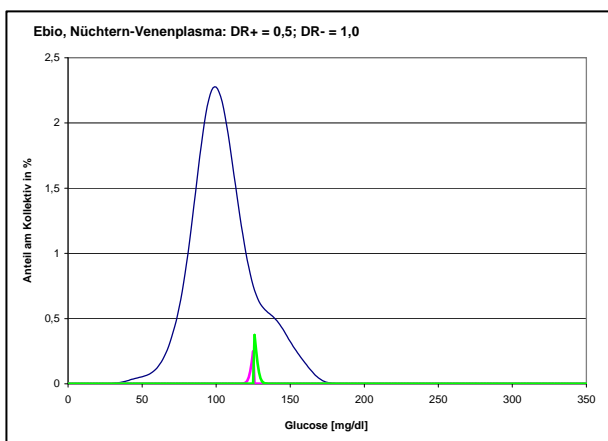
26 b:



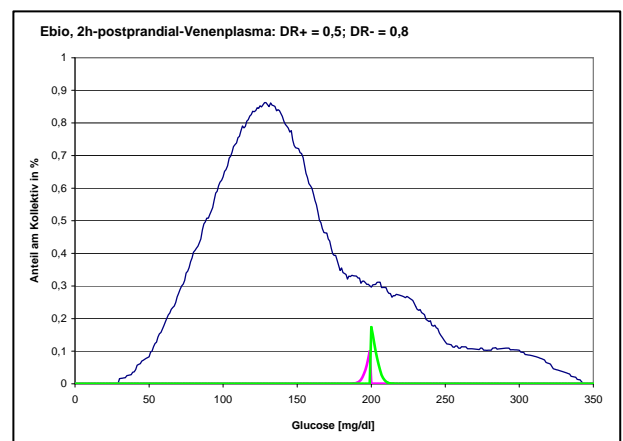
26 e:



26 c:



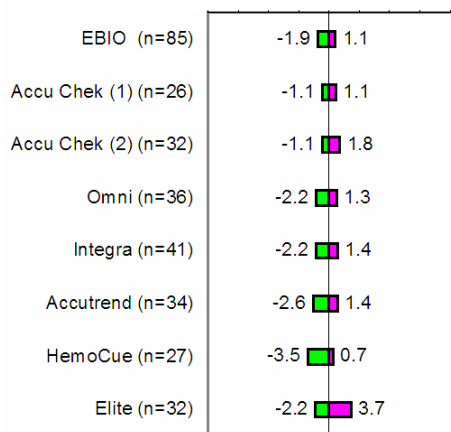
26 f:



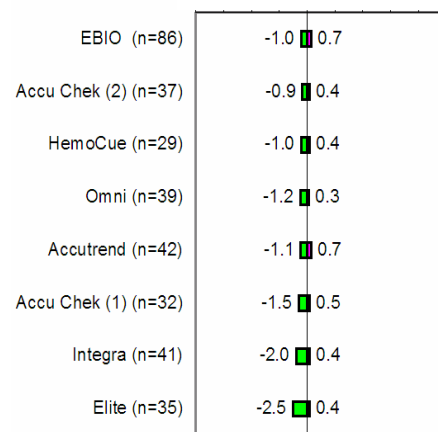
Die Gesamt-DR aller intrasystemischen Vergleiche werden in aufsteigender Reihenfolge in den Abbildungen 27 a – f als Balkendiagramme präsentiert. Dabei entspricht die Längsachse, an der die Balken ausgerichtet sind, der Null-Prozent-DR, positive DR werden durch nach rechts weisende, negative DR durch nach links weisende Balken dargestellt.

Abbildungen 27 a - f: Positive und negative Diskordanzraten (%) der intrasystemischen Vergleiche

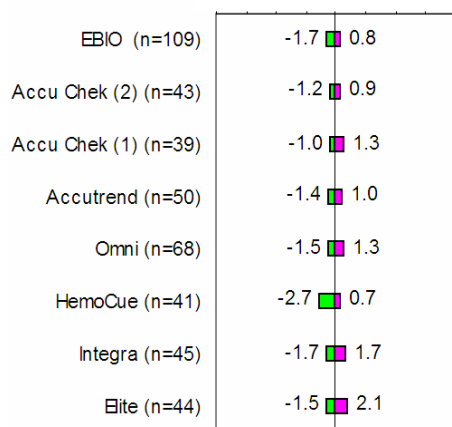
27a) Nüchtern-Kapillarblut



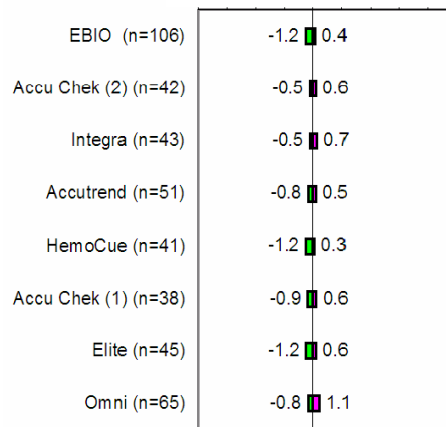
27b) 2h-Postprandial-Kapillarblut



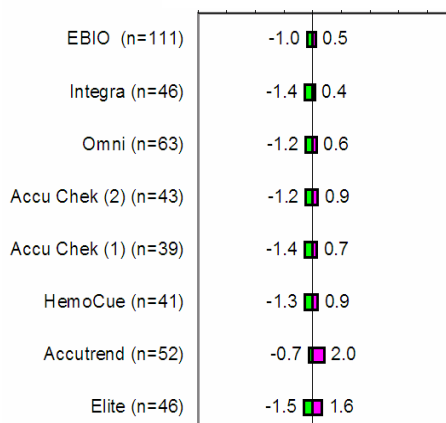
27c) Nüchtern-Venenblut



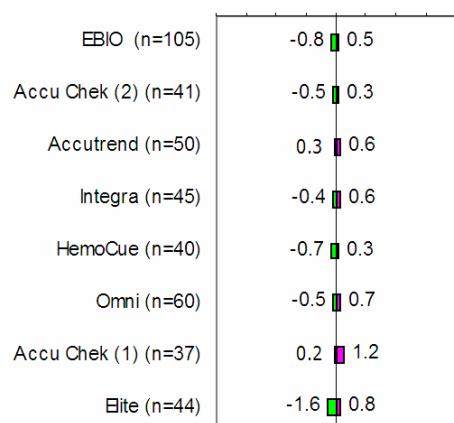
27d) 2h-Postprandial-Venenblut



27e) Nüchtern-Venenplasma



27f) 2h-Postprandial-Venenplasma



4.2.3.2 Diskordanzraten der intersystemischen Vergleiche

Die Tabellen 20 und 21 enthalten die Ergebnisse des Diskordanzraten-Testes der intersystemischen Vergleiche, also die intersystemischen Diskordanzraten. Darin sind die Gesamt-DR, die Anzahl der Wertepaare und die Größe der jeweiligen Populationskollektive (getrennt in Nüchtern- und Postprandial-Zeitpunkt) erfasst. Im oberen Zeilenblock stehen die DR der unkorrigierten, im mittleren Zeilenblock die der Tages-korrigierten und im unteren Zeilenblock die der transformierten intersystemischen Vergleiche.

Tabelle 20: Gesamt-Diskordanzraten (%) der intersystemischen Vergleiche zum Nüchternzeitpunkt

	Matrix	EG	INT	O	AT	HC	E	AC1	AC2	Pop-Koll*
Wertepaare (n)	CBtest / CB	110	45	62	52	44	48	40	45	150
	VBtest / VB	110	46	71	53	45	48	41	45	167
	VPtest / VP	126	47	70	54	45	48	41	45	160
a) Rohdaten	CBtest / CB	110	3,6	37,7	12,9	5,7	15,5	8,5	10,3	
	VBtest / VB	110	4,3	30,7	5,4	7,1	17,9	5,7	8,3	
	VPtest / VP	126	5,1	3,7	5,1	8,5	27,8	6,7	4,7	
b) pTk-korrigiert	CBtest / CB	110	5,6	37,7	12,9	9,4	5,4	32,1	34,9	
	VBtest / VB	110	5,2	31,9	6,5	10,5	6,2	29,8	32,7	
	VPtest / VP	126	3,1	3,8	4,7	10,7	6,7	9,0	13,5	
c) transformierte Rohdaten	CBtest / CB	110	3,1	4,9	6,6	5,2	4,0	4,3	4,0	
	VBtest / VB	110	3,2	5,6	5,3	7,1	6,4	3,8	4,7	
	VPtest / VP	126	5,1	3,7	5,4	5,8	3,5	3,7	4,6	

Tabelle 21: Gesamt-Diskordanzraten (%) der intersystemischen Vergleiche zum 2h-Postprandialzeitpunkt

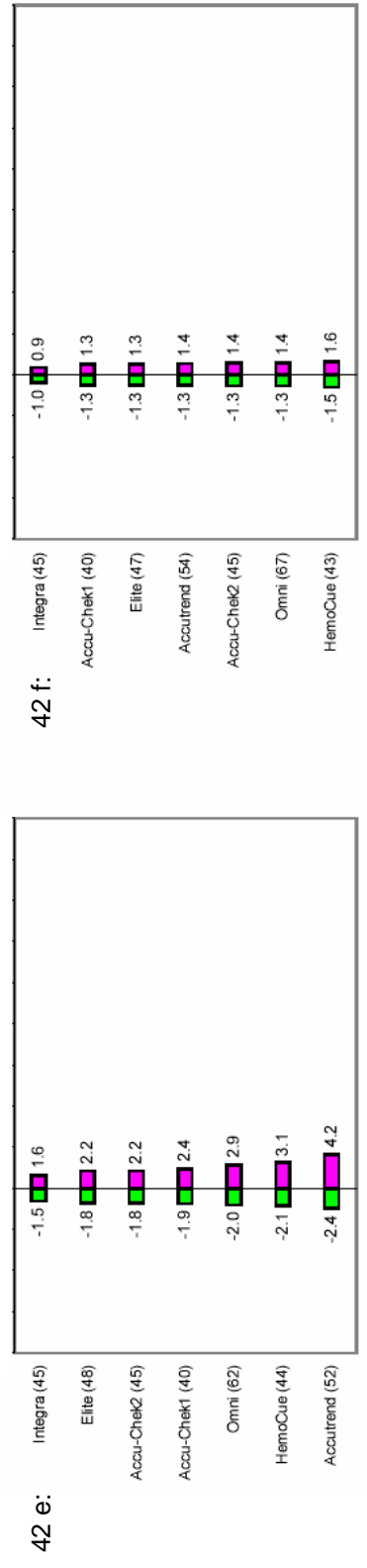
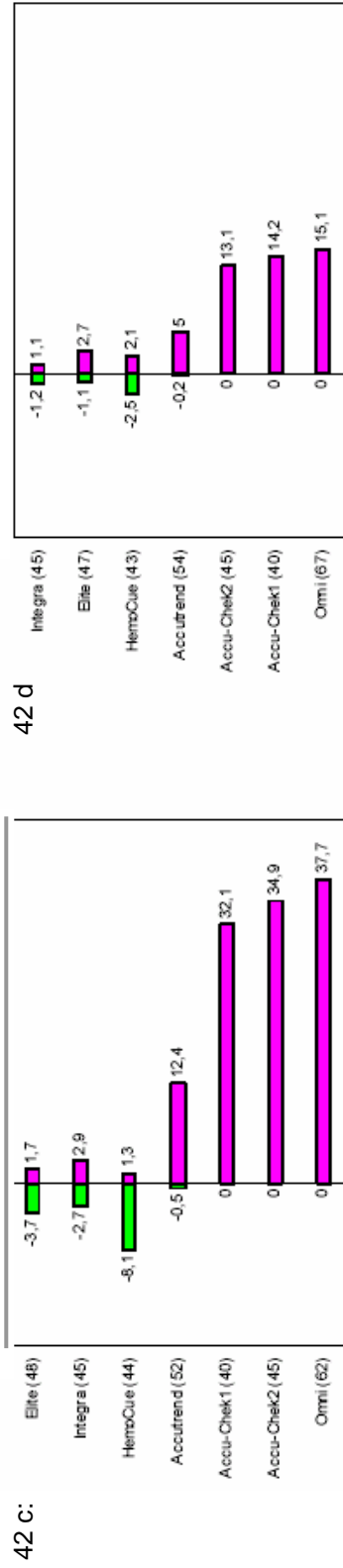
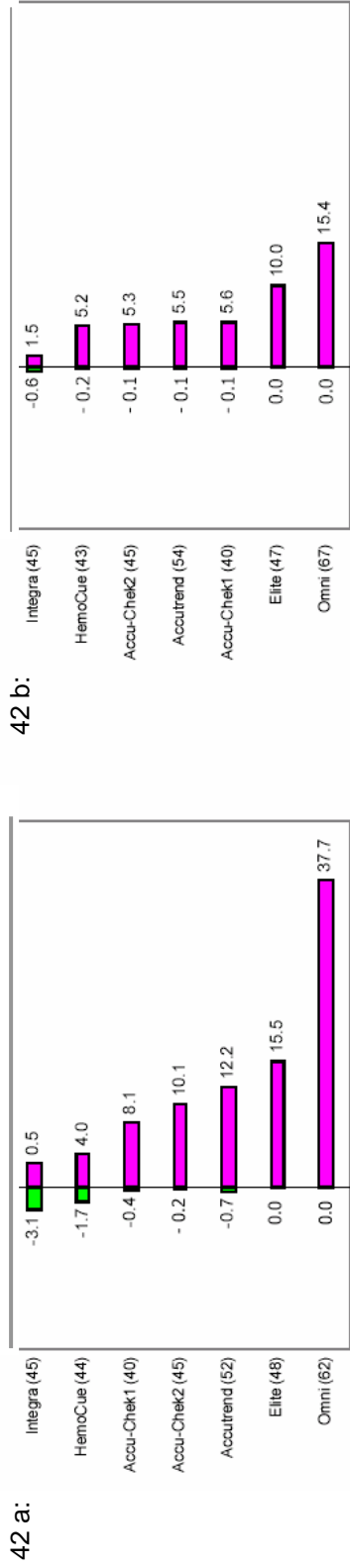
	Matrix	EG	INT	O	AT	HC	E	AC1	AC2	Pop-Koll*
Wertepaare (n)	CBtest / CB	200	45	67	54	43	47	40	45	168
	VBtest / VB	180	44	68	52	43	46	39	43	146
	VPtest / VP	200	46	68	52	43	46	39	43	156
a) Rohdaten	CBtest / CB	200	2,1	15,4	5,6	5,4	10,0	5,7	5,4	
	VBtest / VB	180	1,9	11,0	2,0	2,5	7,8	2,3	2,5	
	VPtest / VP	200	3,1	1,9	2,1	2,4	12,0	7,6	7,6	
b) pTk-korrigiert	CBtest / CB	200	2,3	15,1	5,2	4,6	3,8	14,2	13,1	
	VBtest / VB	180	2,4	10,6	2,1	4,5	3,6	8,8	8,0	
	VPtest / VP	200	1,8	2,0	1,8	6,4	3,7	2,6	2,9	
c) transformierte Rohdaten	CBtest1 / CB1	200	1,9	2,7	2,7	3,1	2,6	2,6	2,7	
	VBtest1 / VB1	180	1,6	2,4	1,8	2,0	2,2	1,8	2,5	
	VPtest1 / VP1	200	2,8	2,0	2,1	2,4	2,7	2,2	3,0	

 Matrix = biologische Matrix, EG = Entscheidungsgrenze [mg/dl],
 ZP = Zeitpunkt, CB = Kapillarblut, VB = Venöses Blut, VP = Venöses Plasma
 INT = Integra, O = Omni, AT = Accutrend, HC = HemoCue, E = Elite, AC = Accu-Chek
 *) Anzahl der Messungen der jeweiligen Matrix des Populationskollektiv
 DR > 5,5% sind A22hellgrau, DR > 11% sind dunkelgrau hervorgehoben

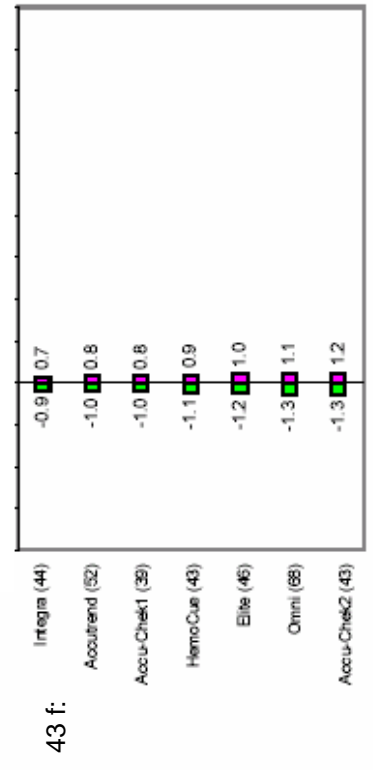
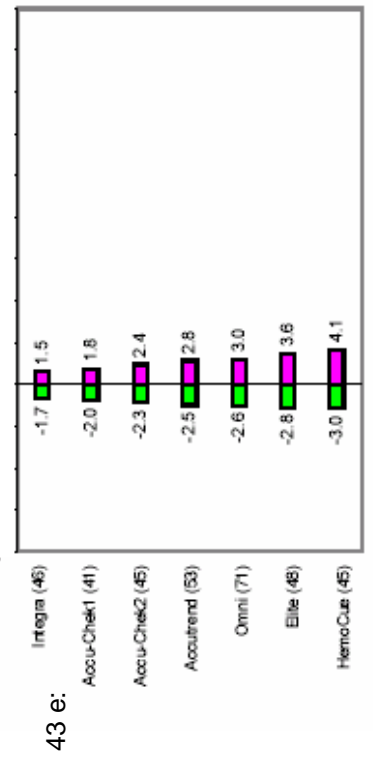
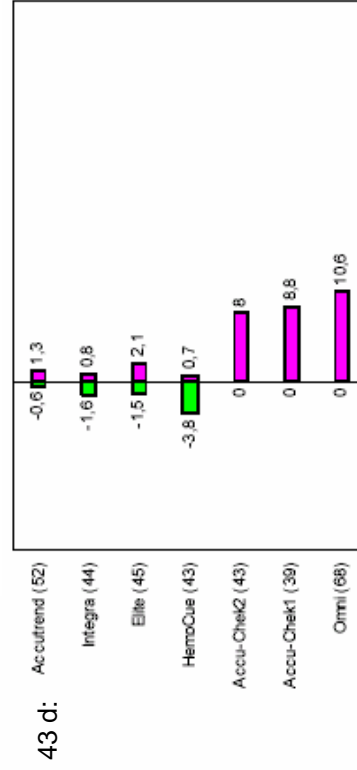
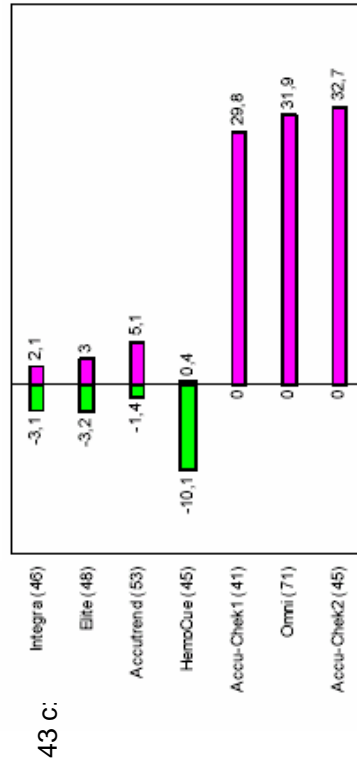
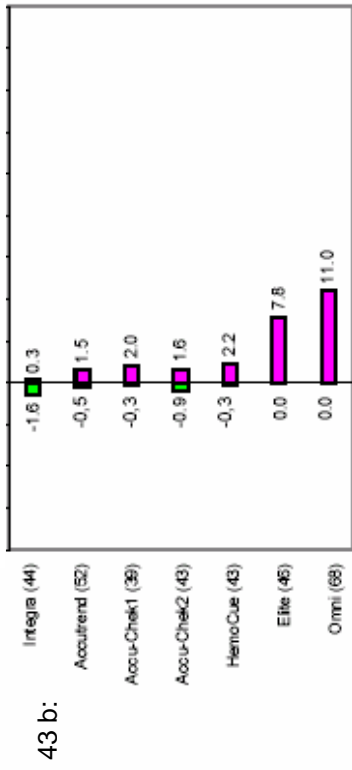
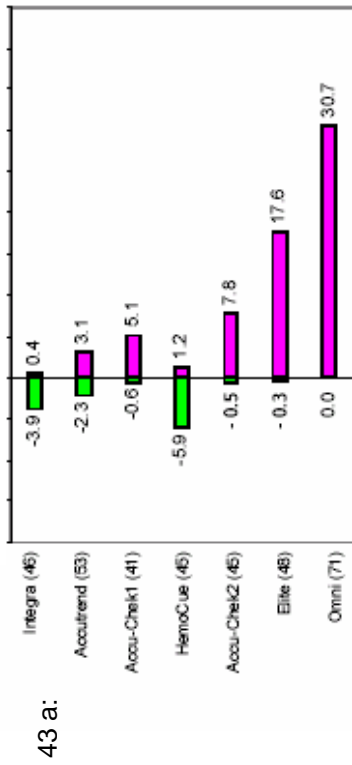
Die DR-Graphiken der unkorrigierten und Tages-korrigierten intersystemischen Vergleiche sind in den Abbildungen 28 – 42 im Anhang hinterlegt. Auf die graphische Darstellung der transformierten intersystemischen Vergleiche wurde verzichtet.

Die Gesamt-DR aller intersystemischen Vergleiche werden in aufsteigender Reihenfolge in den Abbildungen 42 – 43 als Balkendiagramme präsentiert. Dabei entspricht die Längsachse, an der die Balken ausgerichtet sind, der Null-Prozent-DR, positive DR werden durch nach rechts weisende, negative DR durch nach links weisende Balken dargestellt. Links stehen die Balkendiagramme des Nüchtern-Zeitpunktes, rechts die des 2h-postprandial-Zeitpunktes. Jeweils die beiden oberen Balkendiagramme präsentieren die DR der unkorrigierten, die beiden mittleren Balkendiagramme die der Tages-korrigierten und die beiden unteren Balkendiagramme die der transformierten intersystemischen Vergleiche.

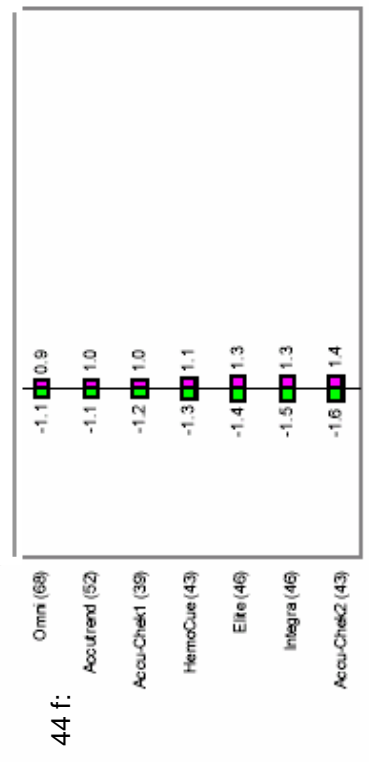
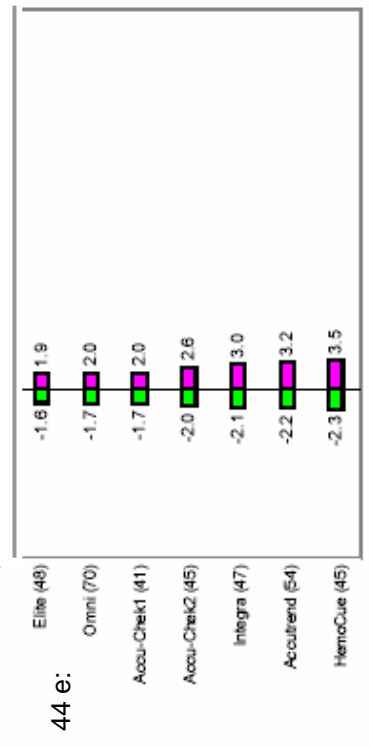
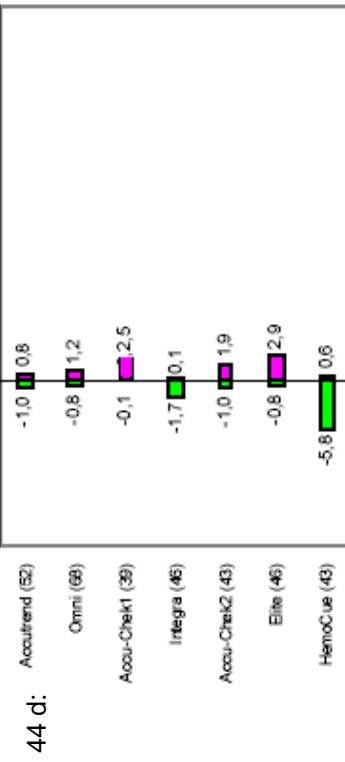
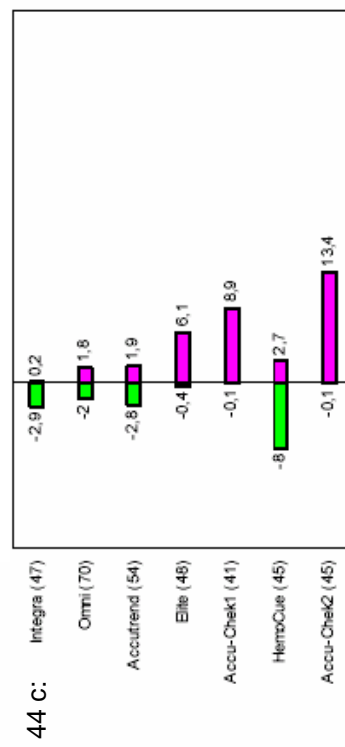
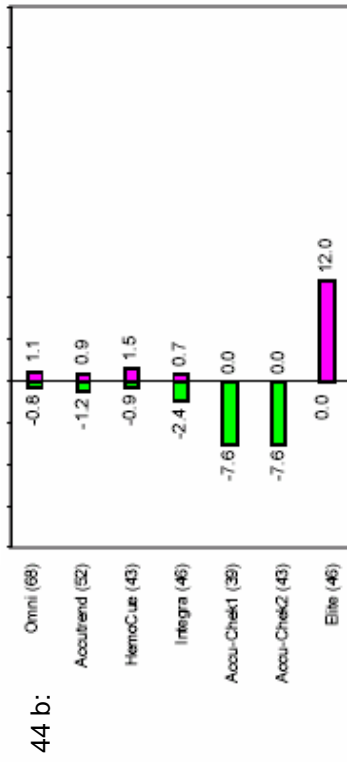
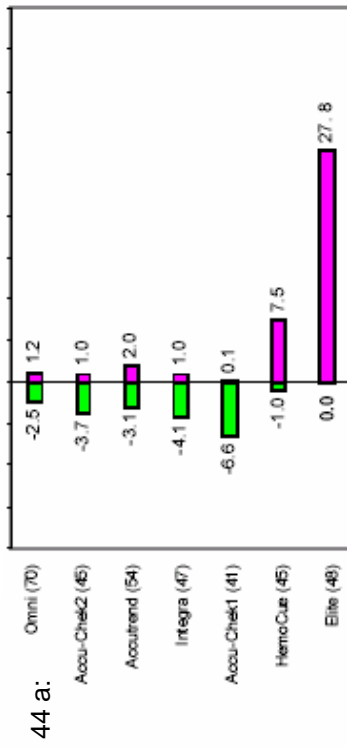
Abbildungen 42 a – f: Positive und negative Diskordanzraten (%) der intersystemischen Vergleiche im Kapillarblut
 (a, b = unkorrigierte Vergleiche; c, d = Tages-korrigierte Vergleiche; e, f = transformierte Vergleiche; a, c, e = nüchtern; b, d, f = 2h-postprandial)



Abbildungen 43 a – f: Positive und negative Diskordanzraten (%) der intersystemischen Vergleiche im venösen Vollblut
 (a, b = unkorrigierte Vergleiche; c, d = Tages-korrigierte Vergleiche; e, f = transformierte Vergleiche; a, c, e = nüchtern; b, d, f = 2h-postprandial)



Abbildungen 44 a – f: Positive und negative Diskordanzraten (%) der intersystemischen Vergleiche im venösen Plasma
 (a, b = unkorrigierte Vergleiche; c, d = Tages-korrigierte Vergleiche; e, f = transformierte Vergleiche; a, c, e = nüchtern; b, d, f = 2h-postprandial)



Die Gesamt-DR zum 2h-postprandialen Zeitpunkt waren kleiner als die zum Nüchternzeitpunkt. Die Gesamt-DR des Vollbluts setzten sich hauptsächlich durch positive DR zusammen, bis auf die von Integra (CB: ausgewogen, VB: DR(neg)), HemoCue (VB: DR(neg)) und Accutrend (VB0: ausgewogen). Die Gesamt-DR des VP setzten sich hauptsächlich durch negative DR zusammen, bis auf die von HemoCue (VP: DR(pos)), Elite (VP: DR(pos)), Accutrend (VP: ausgewogen) und Omni (VP: ausgewogen). Falls nicht näher präzisiert bezieht sich im Folgenden der Begriff Diskordanzrate auf die Gesamt-Diskordanzrate.

Integra erreichte als einziges Vergleichsgerät bei allen Probenarten zu beiden Zeitpunkten DR, die bis maximal 5% lagen (zwischen 1,9% und 4,3%), mit Ausnahme im venösen Plasma, da lag die DR bei 5,1%. Von den POCT-Geräten wurde dieses 5%-Kriterium zum Nüchternzeitpunkt nur noch von Omni und Accu-Chek2 im venösen Plasma erreicht. Zum 2h-postprandialen Zeitpunkt wurde dieses Kriterium von Accutrend und HemoCue im venösen Plasma und Vollblut, von Omni im venösen Plasma und von den beiden Accu-Cheks im venösen Vollblut erreicht. DR bis 10%, was noch als akzeptabel erachtet wurde (Kapitel 3.3.3.3) erzielten HemoCue und Accu-Chek1 in allen Probenarten zu beiden Zeitpunkten, Accutrend und Accu-Chek2 in den venösen Proben und dem 2h-postprandial-Kapillarblut. Omni (Vollblut), Accutrend (Nüchternkapillarblut) und Elite wiesen Gesamt-DR auf, die über 10% lagen (zwischen 10,3% und 37,7%).

Die DR der Tages-korrigierten Testwertvergleiche waren beim Omni und Accutrend quasi unverändert im Vergleich zu den DR der unkorrigierten Testwertvergleiche. Beim Integra überschritten die DR der Nüchtern-Vollblutmessungen nach Tages-Korrektur etwas die 5%-Grenze. Beim HemoCue waren die DR der Tages-korrigierten Testwertvergleiche bis auf eine Ausnahme (CB120) etwas schlechter als die der unkorrigierten Glukosekonzentrationsmesswerte, wobei sich da dann die DR aus negativen DR zusammensetzte. Beim Elite waren die DR der Tages-korrigierten Testwertvergleiche kleiner (also besser) als die der unkorrigierten Glukosekonzentrationsmesswerte, und zwar ausgewogen. Die DR zum Nüchternzeitpunkt lagen zwar nicht mehr über 10%, aber weiterhin noch über der 5%-Grenze, die DR zum 2h-postprandial-Zeitpunkt lagen unterhalb der 5%. Bei den beiden Accu-Cheks waren die DR der Tages-korrigierten Testwertvergleiche deutlich größer (also schlechter) als die der unkorrigierten Testwertvergleiche: bis auf eine Ausnahme (VP120 Verbesserung durch Tages-Korrektur) lagen größere positive DR vor.

Die DR der transformierten Testwertvergleiche waren fast überall besser als die der unkorrigierten Testwertvergleiche. Insbesondere war das Verhältnis der positiven gegenüber

der negativen DR ausgewogen. Die positiven und negativen DR aller Geräte lagen bei allen Probenarten zu beiden Zeitpunkten unter 5%. Die DR aller Geräte lagen bei allen Probenarten zum Nüchternzeitpunkt bis maximal 7,1% und zum 2h-postprandialen Zeitpunkt unter 5%.

5 DISKUSSION

In der vorliegenden Arbeit wurde erstmals der neu entwickelte Diskordanzraten-Test [8-14] für den entscheidungsorientierten Methodenvergleich verschiedener Glukometer (sechs POCT-Glukometer und zwei Laborroutineverfahren) im Hinblick auf die Diagnosestellung eines Typ 2 Diabetes mellitus (Typ 2 DM) angewendet. Der Diskordanzraten-Test berechnet Fehlerraten, sogenannte Diskordanzraten (positive, negative und Gesamt-Diskordanzraten), die dem Anwender Auskunft darüber erteilen, wie groß der Fehler bei seiner diagnostischen Entscheidung sein kann. Er beantwortet also die klinisch relevante Frage, mit wie vielen diskordanten Entscheidungen zu rechnen ist, wenn zwei Messverfahren (hier: Glukometer) in ein und derselben Probe (hier: Blut oder Plasma) eine Messgröße (hier: Glukose) bestimmen und anhand einer Entscheidungsgrenze zwischen zwei Entitäten (in diesem Fall: diabetisch vs. nicht-diabetisch) unterscheiden. Zwei Wahrscheinlichkeiten kommen bei der Diskordanzraten-Berechnung zum Tragen: 1. eine analytische Wahrscheinlichkeit, welche das Fehlerrisiko quantifiziert, das zu erwarten ist, wenn man eine Methode durch eine andere ersetzt, und 2. eine populationsbasierte Wahrscheinlichkeit, die die Verteilung der Messgröße in einem repräsentativen Probandenkollektiv beschreibt.

POCT-Glukometer sind probate Messsysteme für die Therapieüberwachung von Diabetikern. Für die Diagnosestellung wird deren Einsatz jedoch kontrovers diskutiert. Das Zugrundelegen rein analytischer Zuverlässigkeitskriterien (Impräzision, Richtigkeit, mögliche Interferenzen, Vergleich mit etablierten Routineverfahren) ist für die Validierung von POCT-Glukometern unzureichend, da analytische Leistungsmerkmale, die aus konventionellen Methodenvergleichen ermittelt werden, nicht einfach auf klinische Erfordernisse übertragen werden können. Basierend auf klinischer Erfahrung wurden in den letzten Jahren auf dem Gebiet des Diabetes mellitus Ansätze entwickelt, die klinische Kriterien bei der Beurteilung von Labortests in den Vordergrund stellen. Diesen Ansätzen ist gemein, dass sie Prozentbereiche angeben, die eine Interpretation ermöglichen, ob Glukometer für den klinischen Alltag zuverlässige Messwerte liefern oder nicht. Koschinsky *et al.* [6] definierten Akzeptanzbereiche für Vergleiche von Glukometern, die fast identisch mit den „Zonen“ des Error-Grid-Konzepts ($\pm 20\%$) sind, welches von Cox [4] und Clarke [5] entwickelt wurde. Sie postulierten, dass 95% aller Ergebnisse innerhalb eines Akzeptanzbereiches liegen sollen. Dieser Bereich entspricht einer Abweichung des vom Referenzgerät gemessenen „wahren“ Wertes um $\pm 2 s_{y,x}$ (Standardfehler der Residuen) bei Abwesenheit eines Bias. In der Behandlung und der Therapieüberwachung von Diabetikern - die Anwendungsgebiete, auf die sich die erwähnten Konzepte vorrangig beziehen - ist der Bias von Glukometern weniger relevant als die Impräzision. Anders verhält es sich bei der Diagnosestellung eines Diabetes

mellitus Typ 2 und einer gestörten Glukosetoleranz (IGT). Hier sind Impräzision und Bias (Gesamtfehlerrate) gleichermaßen wichtig und erfordern strengere Zielbereiche. Diese strengere Herangehensweise kann durch eine geringere Variabilität der Glukosekonzentration begründet werden. Die präanalytische Phase kann wesentlich einfacher bei Personen standardisiert werden, die für die Diagnose eines Typ 2 DM selektiert werden, als bei Diabetikern, deren Glukose aus therapeutischen Gründen kontrolliert werden muss. Es handelt sich bei diesen Personen in der Regel um mehr oder weniger „gesunde“ Patienten, bei denen elektiv die Stoffwechselsituation vorzugsweise im Rahmen von oralen Glukosetoleranztests abgeklärt werden. Faktoren, die die Glukosekonzentration beeinflussen können, fallen weg. Beispielsweise schwankt der Hämatokrit deutlich geringer als bei klinisch kranken Personen; Hyperlipoproteinämie als weitere mögliche Fehlerquelle entfällt ebenfalls als Störfaktor, da im Nüchternblut die Triglycerid-Konzentration im Blut vernachlässigbar ist.

Vor diesem Hintergrund wurde der Diskordanzraten-Test entwickelt, der sich – wie Spezifität, Sensitivität und Effizienz – an diagnostischen Kriterien orientiert. Zu der klassischen Beurteilung der diagnostischen Zuverlässigkeit eines Labortestes in Methodenvergleichen wird häufig die Non-Effizienz (1 minus diagnostische Effizienz) herangezogen, die der reinen ausgezählten Rate an falsch positiven und falsch negativen Messungen entspricht, welche der Funktion folgt: $\text{Non-Effizienz} = \text{falsche Messergebnisse} / \text{alle Messergebnisse}$ (siehe Tabelle 11 im Anhang). Wird die Non-Effizienz aus einer kleinen Stichprobe berechnet, dann kann sie leicht überschätzt aber auch unterschätzt werden. Eine ausreichend große Fallzahl ($n > 100$), um nach klassischer Herangehensweise robuste Messergebnisse zu erhalten, kann sich bei Methodenvergleichen als zu aufwendig erweisen. Der Diskordanzratentest ermöglicht aufgrund auf Wahrscheinlichkeiten beruhender Annahmen, aus kleinen Stichproben ($n < 50$) verlässliche Ergebnisse zu berechnen. Das statistische Konzept der diagnostischen Non-Effizienz wird beim Diskordanzraten-Test erweitert, indem der Effekt der Variabilität um eine Entscheidungsgrenze und die Verteilung der Messwerte in der Patientenpopulation beim Diskordanzratentest berücksichtigt werden.

In der vorliegenden Arbeit wurde also zunächst die analytische Zuverlässigkeit von sechs POCT-Glukometern und zwei Laborroutinemessgeräten anhand von Messungen in Standard- und Kontroll-Lösungen, als auch in biologischen Proben überprüft. Mittels des Diskordanzraten-Tests wurde im Anschluss daran die klinische Relevanz der zwischen den Testgeräten und dem Referenzgerät sich darstellenden Differenzen hinsichtlich des diagnostischen Entscheidungsprozesses ermittelt.

Die Überprüfung der analytischen Zuverlässigkeit in Standard- und Kontroll-Lösungen erfolgte in selbst angesetzten primären wässrigen Glukosestandards ohne und mit Albumin-Zusatz (Glu-NIST, Glu-Alb), deren Glukosegehalt bei 200 mg/dl lag, und in zwei handelsfertigen proteinhaltigen lyophilisierten Kontroll-Lösungen (Val A und Val N), deren Glukosegehalt bei 208 mg/dl respektive 93,9 mg/dl lagen. Zur Validierung, Qualitätssicherung und Kalibration von Messmethoden ist es notwendig, Messungen in Lösungen mit vordefinierten Konzentrationen der zu untersuchenden Messgröße durchzuführen [55-58]. Wichtig dabei ist, dass Lösungen mit unterschiedlich konzentrierter Messgröße getestet werden. Ebenso sinnvoll ist es, unterschiedlich zusammengesetzte Lösungen zu analysieren, um das Ausmaß möglicher beeinflussender Störeffekte zu erkennen. Rein wässrige Ansätze sind weitestgehend frei von Störeffekten. Wässrige Ansätze mit Albuminzusatz ermöglichen dagegen, den Einfluss von Protein auf die Messergebnisse abzuschätzen. Kontroll-Lösungen wie Val A und Val N enthalten viele verschiedene Messgrößen, und entsprechen so in ihrer Zusammensetzung eher einer biologischen Probe. Üblicherweise wird jedes Messgerät mit gerätespezifischen Kalibrier/Kontroll-Lösungen auf den Markt gebracht. Solche gerätespezifischen Lösungen haben den Nachteil, dass sie nur für das jeweilige Gerät zu nutzen und oft teuer sind [55]. Ein Vorteil von selbst angesetzten primären Standardlösungen, z.B. mit reiner Glukose des National Institute of Standards and Technology (NIST), ist, dass sie geräteunabhängig eingesetzt werden können. Laborroutinegeräte werden häufig durch solche rein wässrigen Standardlösungen überprüft; bei den POCT-Glukometern hingegen ist der Einsatz solcher Standardlösungen oft aus analytischen Gründen nicht möglich [56-59]. Bei Kontroll-Lösungen, deren Qualität und Zusammensetzung nicht immer nachzuvollziehen ist, können unvorhersehbare Einflüsse die Messergebnisse verfälschen und deren Interpretation erschweren [60].

Von den acht Messsystemen haben Ebio, Integra, Omni und Accutrend die Standard- und Kontroll-Lösungen analytisch zuverlässig gemessen, die Messwerte lagen nahe bei den Sollwerten. Die Messwerte des Elite und des HemoCue lagen oberhalb, die der beiden Accu-Cheks (mit Ausnahme der Val A-Messungen) unterhalb der Sollwerte (siehe Tabelle 12 und Abbildungen 10 in Kapitel 4.1.1). Daraus ist ableitbar, dass die Messqualität von Ebio, Integra, Omni und Accutrend durch einen einfach ansetzbaren Glukosestandard und durch geräteunabhängige Kontroll-Lösungen kontrollier- bzw. überprüfbar ist. Die Sollwert-Abweichungen, wie sie bei Elite, HemoCue und Accu-Chek beobachtet wurden, können für Kalibrationszwecke genutzt werden, sofern das jeweilige Gerät mit gleicher Abweichungsrichtung und gleichem Abweichungsgrad in biologischen Proben misst. Wenn also Elite und HemoCue nicht nur die Standard- und Kontroll-Lösungen, sondern auch die

biologischen Proben zu hoch messen, dann können die Messwerte korrigiert werden. Vice versa gilt das für Accu-Chek.

Die Überprüfung der analytischen Zuverlässigkeit in biologischen Proben erfolgte in Blut- und Plasmaproben, die zu zwei verschiedenen Entnahmezeitpunkten im Rahmen von 78 oralen Glukosetoleranztests, die der Abklärung einer diabetischen Stoffwechsellage dienten, gemessen wurden. Die intrasystemischen Vergleiche zeigten erwartungsgemäß eine exakte Korrelation, woraus abgeleitet werden kann, dass die Messungen reproduzierbar waren. Bei den intersystemischen Vergleichen (siehe Tabellen 16 und 17 in Kapitel 4.2.2) fiel auf, dass nur Integra bei allen Messungen mit dem Ebio übereinstimmte. Elite lieferte durchweg zu hohe Messwerte. Omni maß nur im venösen Plasma präzise. Das venöse Plasma wurde ansonsten nur noch von Accutrend verlässlich gemessen; Accu-Chek mit zwar geringen Mittelwert-Abweichungen von <6% war schon nicht mehr zuverlässig. Das Kapillarblut wurde nur von HemoCue analytisch präzise bestimmt. Accutrend und Accu-Chek, deren Mittelwert-Abweichungen unter 7% lagen, konnten infolgedessen nicht mehr als analytisch zuverlässig gelten. Anders verhielt es sich bei den Messungen von venösem Vollblut. Da waren die Messwerte von HemoCue, Accutrend und Accu-Chek vergleichbar mit denen des Ebio. Die intersystemischen Methodenvergleiche wurden mit den Tages-korrigierten Messwerten wiederholt, um insbesondere der Frage nachzugehen, ob die POCT-Glukometer in den Probenarten, in denen sie analytisch nicht zuverlässig messen, durch den Bezug auf einen wässrigen Standard hinsichtlich ihrer analytischen Leistungsfähigkeit verbessert werden können. Ob also durch eine tägliche Korrektur mittels eines Quotienten aus Ist- und Sollwert eines wässrigen Glukosestandards die Messwerte dahingehend korrigiert bzw. kalibriert werden, dass sie den Ebio-Ergebnissen vergleichbar sind (Tages-Korrektur durch Bezug auf den wässrigen Glu-NIST-Standard = Rückführbarkeit). Bei Elite konnte durch die Tages-Korrektur eine analytische Zuverlässigkeit erreicht werden, da Elite sowohl die verschiedenen Proben als auch den wässrigen Glukosestandard in gleichem Ausmaß zu hoch gemessen hatte, so dass Elite demzufolge auf eine primäre Standardlösung rückgeführt werden konnte. Die Messungen der anderen Glukometer wurden durch die Tages-Korrektur nicht geändert und waren somit nicht auf einen primären Standard rückführbar.

Für alle sich ergebenden intra- und intersystemischen Methodenvergleiche wurden die positiven, negativen und Gesamt-Diskordanzraten in Bezug auf die WHO-definierten Entscheidungsgrenzen, die eine diabetische von einer nicht-diabetischen Stoffwechsellage trennt [42;47-49], berechnet. Berechnungen erfolgten für die intrasystemischen Vergleiche und für die unkorrigierten, Tages-korrigierten und transformierten (Bias-eliminierten)

intersystemischen Vergleiche. Die analytischen Wahrscheinlichkeiten wurden aus den Ergebnissen der Passing-Bablok-Regressionsanalysen abgeleitet. Die Populationswahrscheinlichkeiten wurden auf Basis von 168 oralen Glukosetoleranztests berechnet, die von 1994 bis 2001 im Zentrallabor der Klinik durchgeführt worden waren. Als Zielgröße für die Genauigkeit der Geräte wurde eine Gesamt-Diskordanzrate von $\leq 5\%$ gewählt. Eine Diskordanzrate von 5% bedeutet, dass mindestens 95% des Studienkollektivs demselben diagnostischen Kollektiv zugeordnet werden. Geräte, deren Gesamt-Diskordanzraten unter 10% lagen, deren positive und negative Diskordanzraten allerdings nicht größer als 5% waren, wurden als akzeptabel eingestuft. Geräte deren Diskordanzraten darüber lagen, galten als inakzeptabel.

Es ergaben sich für die 2h-postprandialen Messungen niedrigere Diskordanzraten, als für die Nüchtern-Messungen. Dies kann dadurch bedingt sein, dass sich die Messwerte zwei Stunden postprandial bei Patienten mit diabetischer und nicht-diabetischer Stoffwechsellage stärker unterscheiden. In beiden Fällen würde eine geringere analytische Wahrscheinlichkeit für das Auftreten von Diskordanzen und letztendlich eine kleinere Diskordanzrate resultieren. Die Diskordanzraten der intrasystemischen Vergleiche lagen (fast) alle erwartungsgemäß deutlich unter 5%, also eindeutig im präzisen Bereich. Nur Elite wies im Nüchternkapillarblut eine Diskordanzrate von 5,9% auf.

Bei den intersystemischen Diskordanzraten (wenn nicht anders spezifiziert, sind im folgenden Text die Gesamt-Diskordanzraten gemeint) der unkorrigierten Methodenvergleiche (siehe Tabellen 20 und 21 und Abbildungen 42 bis 44 in Kapitel 4.2.3.2) waren die Ergebnisse pro Gerät betrachtet folgendermaßen: Integra führte als einziges Gerät in allen biologischen Probenarten zu Diskordanzraten, die unter 5% lagen (wobei die DR des Nüchtern-Venenplasmas bei 5,1% lag). Das war zu erwarten, da es sich bei Integra ebenso wie beim Referenzgerät Ebio um ein Routinelaborgerät handelt, das vergleichbare Qualitätskriterien wie das Ebiogerät erfüllt. Integra diente quasi als Kontrolle des neuen Verfahrens. Die größten Diskordanzraten (>30%) wurden für die Vollblutmessungen des Omni ermittelt (Nüchternzeitpunkt). Bei den Messungen im venösen Plasma hingegen zeigte Omni mit <4% die niedrigsten Diskordanzraten. Omni ist allerdings auch primär für die Messung im Plasma und nicht im Vollblut konzipiert. Die zellulären Bestandteile stellen einen erheblichen Störfaktor bei der Bestimmung der Glukoseaktivität in der wässrigen Phase dar. Dieses Prinzip wird jedoch von Omni für die Bestimmung der Glukosekonzentration verwendet. Bei keinem der portablen POCT-Geräte wurden Diskordanzraten festgestellt, die zu beiden Zeitpunkten des oralen Glukosetoleranztestes unter 5% lagen. Mit Ausnahme des HemoCue lagen die Gesamt-DR der portablen Geräte im kapillären Vollblut höher als in den

anderen Probenarten. Dies ist bemerkenswert, da sie für Messungen im Kapillarblut konzipiert sind. HemoCue erreichte in allen Probenarten zu beiden Zeitpunkten immerhin noch Gesamt-Diskordanzraten von unter 10%. Auffällig ist, dass bei HemoCue die Diskordanzrate des kapillären Vollbluts und des venösen Plasmas hauptsächlich durch positive Diskordanzraten, während die des venösen Vollblutes hauptsächlich durch negative Diskordanzraten bedingt sind. Dafür gab es keine plausible Erklärung. Die Diskordanzraten der beiden Accu-Chek-Geräte lagen zum Nüchternzeitpunkt zwischen 4,7% und 10,3%, und zum 2h-postprandial-Zeitpunkt zwischen 2,3% und 7,6%. Auffällig war, dass hier die Diskordanzraten des Vollbluts hauptsächlich durch positive Diskordanzraten, während die des Plasmas hauptsächlich durch negative Diskordanzraten bedingt waren. Das könnte durch die zellulären Bestandteile als Störgröße erklärt werden, die tendenziell zu höheren Glukosemesswerten führen. Elite wies als einziges Gerät in allen biologischen Probenarten Diskordanzraten auf, die über 5% lagen, zum Nüchternzeitpunkt sogar durchgängig über 10%, hauptsächlich durch positive Diskordanzraten bedingt. Elite kann unkorrigiert nicht als verlässliches Glukometer akzeptiert werden, insbesondere nicht, wenn eine diagnostische Entscheidung aus den Messungen abgeleitet werden soll.

Die Diskordanzraten der Tages-korrigierten Methodenvergleiche des Integra, Omni und Accutrend blieben weitgehend unverändert im Vergleich zu denen der unkorrigierten Methodenvergleiche, wobei die DR des Integra beim Nüchternvollblut allerdings knapp die 5%-Grenze überschritt (CB0 5,6%, VP0 5,2%), beim Nüchternvenenplasma aber ausnahmslos in den 5%-Bereich korrigiert wurde. Beim HemoCue waren die Diskordanzraten der Tages-korrigierten Methodenvergleiche bis auf eine Ausnahme (CB120 6,4%) etwas schlechter, so dass die DR des venösen Vollbluts und Plasmas dann auch schon nicht mehr im akzeptablen Bereich (VB0 10,5%, VP0 10,7%) lagen. Beim Elite waren die Diskordanzraten der Tages-korrigierten Methodenvergleiche deutlich besser, bei den beiden Accu-Cheks waren sie bis auf eine Ausnahme (VP120) deutlich schlechter als die der unkorrigierten Methodenvergleiche.

Nach Bias-Ausschluss konnten alle Glukometer mindestens als akzeptabel, wenn nicht sogar als präzise eingestuft werden. Die Diskordanzraten der transformierten Methodenvergleiche lagen immerhin alle unter 10% (beim 2h-postprandial-Zeitpunkt unter 5,0%), wenn auch die Gesamt-DR zum Nüchternzeitpunkt von Accutrend und HemoCue bei allen Probenarten und von Omni und Elite beim venösen Vollblut noch über 5% lagen, weshalb diese Geräte auch nicht als gleichwertig zum etablierten Ebio eingestuft werden können. Integra blieb durch die Transformation weiterhin vergleichbar gut wie Ebio, die Vollblutvergleiche des Omni wurden deutlich gebessert und waren nach Transformation mit

dem Ebio vergleichbar. Die Diskordanzraten des Elite wurden ebenfalls deutlich gebessert und lagen nach Transformation (fast) allesamt unter 5% (bis auf 6,4% im venösen Vollblut, Nüchternzeitpunkt). HemoCue lag unverändert im gerade noch akzeptablen Bereich. Die Diskordanzrate des Accutrend im Kapillarblut (Nüchternzeitpunkt) war nach Transformation zwar gebessert (auf 6,6% reduziert), lag demzufolge nicht mehr im nicht-akzeptablen aber auch noch nicht im präzisen Bereich. Die Diskordanzraten der beiden Accu-Cheks waren nach Transformation alle mit Ebio vergleichbar.

Daraus kann gefolgert werden, dass Integra in allen Probenarten, Omni im venösen Plasma und Accutrend im venösen Vollblut und Plasma die Gukosekonzentration vergleichbar gut wie Ebio messen und diagnostische Entscheidungen fast identisch getroffen werden. Bei den Messungen mit HemoCue und Accuchek treten gehäuft Fehlentscheidungen auf, bei Elite ist die Fehlentscheidungsrate nicht mehr tolerabel. Die Tages-Korrektur (die Rückführung mittels NIST-Standard), bewirkte nur bei Elite eine Verbesserung.

Als Fazit geht daraus hervor, dass beide Laborroutinemessgeräte gleich gut für diagnostische Entscheidungen herangezogen werden können. Omni eignet sich für diagnostische Zwecke, solange die Messungen im venösen Plasma erfolgen. Die portablen POCT-Glukometer hingegen, die für Messungen im Kapillarblut konzipiert sind, eignen sich nicht für diagnostische Zwecke. Selbst nach Bias-Ausschluss, wodurch die Fehlentscheidungen insgesamt deutlich reduziert werden, sind nicht alle getesteten POCT-Glukometer gleichermaßen für den diagnostischen Einsatz geeignet, da immer noch Diskordanzraten über 5,5% auftreten: Accutrend im CB0, HemoCue im VB0 und VP0, Elite und Omni im VB0. Anders ausgedrückt eignen sich von den POCT-Glukometern nach Bias-Elimination Accu-Chek in allen Probenarten, Accutrend im venösen Vollblut und Plasma, Omni und Elite im Kapillarblut und venösen Plasma und HemoCue im Kapillarblut für diagnostische Zwecke.

Als Konsequenz ergibt sich daraus für den klinischen Alltag, dass es nicht ratsam ist, POCT-Glukometer zur Diagnostik eines Typ 2 Diabetes mellitus zu verwenden. Aus den Ergebnissen der linearen Regressionsanalyse könnte man zwar ableiten, dass HemoCue (bei Messungen des Vollbluts), Accutrend (bei Messungen des venösen Blutes und Plasmas), Accu-Chek (bei Messungen des venösen Vollbluts) und Omni (bei Messungen des venösen Plasmas) für die Diagnostik geeignet seien. Allerdings relativiert sich diese Aussage durch die Ergebnisse des Diskordanzraten-Tests, wo gezeigt wurde, dass im Prinzip nur Omni verlässliche Ergebnisse bei den Messungen des venösen Plasmas erzielte. Die niedrigen Diskordanzraten, die nach Transformation erreicht wurden, lassen darauf

schließen, dass POCT-Glukometer vergleichbare Glukosekonzentrations-Ergebnisse wie die etablierter Laborroutineanalysegeräte messen könnten, vorausgesetzt sie sind korrekt kalibriert.

Die aus dieser Studie gezogene Erkenntnis, POCT-Glukometer nicht zur Diagnostik eines Typ 2 Diabetes mellitus zu verwenden, ist im Einklang mit der Empfehlung der Deutschen Diabetes Gesellschaft (DGD) [49]. In der evidenzbasierten Leitlinie Diabetes mellitus der DGD wird ausdrücklich darauf hingewiesen, dass die Diagnose eines Diabetes mellitus nur mit Glukosewerten gestellt werden sollten, die mit einer qualitätskontrollierten Labormethode gemessen wurden. Dort wird ausdrücklich darauf verwiesen, dass Geräte zur Blutzuckerselbstmessung für diagnostische Zwecke ungeeignet seien.

6 Zusammenfassung

In der vorliegenden Arbeit wurde erstmals der Diskordanzraten-Test [8-14] zur Beurteilung der diagnostischen Effizienz verschiedener Glukosemessgeräte angewendet. Dieser Test ist ein neuartiges statistisches Konzept, welches analytische und diagnostische Kriterien miteinander kombiniert. Es wurden Vergleiche sechs verschiedener Point-of-Care-Testing (POCT) Glukometer und zwei Laborroutinemessgeräten (Integra, Ebio) untereinander und gegenüber Ebio durchgeführt. Als Zielgröße wurde eine Gesamt-Diskordanzrate von <5% zwischen den unterschiedlichen Glukose-bestimmenden Systemen festgelegt. Eine Diskordanzrate von 5% bedeutet, dass mindestens 95% des Studienkollektivs demselben diagnostischen Kollektiv zugeordnet werden. Diese Rate hängt von der Verteilung der Glukosekonzentration im Referenzkollektiv und von den analytischen Unterschieden der Messmethoden beim Methodenvergleich ab. Dieser Diskordanzraten-Test überträgt Daten, die man aus einem Methodenvergleich mit kleiner Fallzahl erhebt, auf eine größere Population, die an sich notwendig wäre, um für diagnostische Zwecke genutzt werden zu können. Eine annähernde konstante Verteilung der relativen Differenz $f(x) = (x - y) \times 100$, also der normalisierten Differenzen, wurden für das Daten-Intervall angenommen. Das 5%-Kriterium wurde von etablierten Laborroutinemessverfahren (Integra vs. Ebio) in allen drei Probenarten erreicht. Mit den POCT-Glukometern wurde das 5%-Kriterium erst nach Bias-Ausschluss (Transformation) erreicht, wobei einige Geräte immer noch Diskordanzraten bis zu 7,1% aufwiesen.

Ein ideales POCT-System sollte mittels einer primären Standardlösung oder mittels Kontrollmaterials auf die Referenzmethodenwerte des jeweiligen Labors rückzuführen sein. Das war bei dem Elite-Gerät der Fall, da sich hier die Tages-korrigierten Messergebnisse in den Akzeptanzbereich verschoben, wohingegen die unkorrigierten Messwerte deutlich zu hoch lagen. Darüberhinaus sollte ein ideales POCT-System akzeptable Messwerte nach Bias-Ausschluss (Transformation) erzielen. Die Accu-Chek-Geräte waren die einzigen Glukometer, die nach Bias-Ausschluss in allen Probenarten zu beiden Entnahmezeitpunkten zu Diskordanzraten führten, die unter 5% lagen (Tabellen 20 und 21, Abbildungen 42 – 44).

POCT-Glukometer haben heutzutage in der Therapieüberwachung von Diabetikern ein breite Akzeptanz gefunden, der diagnostische Einsatz wird aber kritisch hinterfragt. Die WHO rät davon ab, POCT-Glukometer zur Diagnose eines Typ 2 Diabetes mellitus anzuwenden, was durch die Ergebnisse der vorliegenden Arbeit unterstützt wird. Unter den in dieser Arbeit angewandten Bedingungen sollte insbesondere eine gestörte Glukosetoleranz nur durch ein etabliertes Labormessverfahren diagnostiziert werden. Einschränkend ist anzumerken, dass

alle hier getesteten POCT-Glukometer nach angemessener Kalibration annähernd so gut wie die etablierten Laborverfahren geeignet waren, einen Typ 2 Diabetes mellitus aufzudecken. Demzufolge wäre es von Seiten der Hersteller sinnvoll, das Kalibrationskonzept zu überdenken.

7 LITERATURVERZEICHNIS

- (1) Bland JM, Altman DG. Statistical methods for assessing agreement between two methods of clinical measurement. *Lancet* 1986; 1(8476):307-310.
- (2) Hollis S. Analysis of method comparison studies. *Ann Clin Biochem* 1996; 33(Pt 1):1-4.
- (3) Fraser CG, Petersen PH. Analytical performance characteristics should be judged against objective quality specifications. *Clin Chem* 1999; 45:490-493.
- (4) Cox D, Clarke WL, Gonder-Frederick LA, Pohl SL, Hoover C, Snyder A. Accuracy of perceiving blood glucose in IDDM. *Diabetes Care* 1985; 8:529-536.
- (5) Clarke WL, Cox D, Gonder-Frederick LA, Carter W, Pohl SL. Evaluating clinical accuracy of systems for self-monitoring of blood glucose. *Diabetes Care* 1987; 10(5):622-628.
- (6) Koschinsky T, Dannehl K, Gries FA. New approach to technical and clinical evaluation of devices for self-monitoring of blood glucose. *Diabetes Care* 1988; 11(8):619-629.
- (7) Boyd JC, Bruns DE. Quality specifications for glucose meters: assessment by simulation modeling of errors in insulin dose. *Clin Chem* 2001; 47:209-214.
- (8) Haeckel R, Püntmann I. Validation of measurement procedures. *In vitro veritas* 2001; 2(art 20): <http://www.acclc.cat/invitroveritas/vol2/art20.html>.
- (9) Püntmann I, Wosniok W, Haeckel R. Application of the diagnostic discordance rate in comparing various POCT glucometers for the diagnosis of type 2 diabetes. *Clin Chem Lab Med* 2001; 39(9).
- (10) Püntmann I, Wosniok W, Haeckel R. A new test for the diagnostic relevance in method comparisons (diagnostic discordance rate). *Clin Chem Lab Med* 2001; 39(9):A60.
- (11) Püntmann I, Wosniok W, Haeckel R. Ein neues Maß für die diagnostische Relevanz beim Methodenvergleich. In: Lison AE, Diehl HA, editors. *Medizinische Forschung und Gesundheitswissenschaften in Bremen*. Universitätsdruckerei Bremen, 2001.
- (12) Haeckel R, Wosniok W, Puentmann I. Discordance rate, a new concept for combining diagnostic decisions with analytical performance characteristics. 1. Application in method or sample system comparisons and in defining decision limits. *Clin Chem Lab Med* 2003; 41(3):347-355.
- (13) Puntmann I, Wosniok W, Haeckel R. Comparison of several point-of-care testing (POCT) glucometers with an established laboratory procedure for the diagnosis of type 2 diabetes using the discordance rate. A new statistical approach. *Clin Chem Lab Med* 2003; 41(6):809-820.
- (14) Haeckel R, Wosniok W. The discordance rate, a new concept for combining diagnostic decisions with analytical performance characteristics. 2. Defining analytical goals applied to the diagnosis of type 2 diabetes by blood glucose concentrations. *Clin Chem Lab Med* 2004; 42(2):198-203.
- (15) Eppendorf AG. *Ebio® plus - Bedienungsanleitung*. Eppendorf Dokument 2000;84 Seiten.
- (16) Kohl H. *Qualitätsmanagement im Labor: Praxisleitfaden für Industrie, Forschung, Handel und Gewerbe*. Springer, 1996.

- (17) Edelhäuser R, Eßer F, Fischbach R, da Fonseca-Wollheim F, Haeckel R, Kindler M, Köppen B, Kruse-Jarres J, Schellenberg I, Schoerner C, Weidemann G. Leitfaden zur Umsetzung der Norm DIN EN 45001 und des ISO Guide 25. In: Sektorkomitee Medizinische Laboratorien ZLG, editor. Medizinisches Labor - Qualitätsmanagement und Akkreditierung. Wissenschaftliche Verlagsgesellschaft mbH, 1997: 1-27.
- (18) Bundesärztekammer. Richtlinie der Bundesärztekammer zur Qualitätssicherung quantitativer laboratoriumsmedizinischer Untersuchungen. Deutsches Ärzteblatt 2001; 98(42):A-2747.
- (19) Bundesärztekammer. Richtlinie der Bundesärztekammer zur Qualitätssicherung quantitativer laboratoriumsmedizinischer Untersuchungen. Bundesärztekammer Dokument 2002;25 Seiten.
- (20) European Committee for Clinical Laboratory Standards. Guidelines for the evaluation of analyzers in clinical chemistry, 3rd draft. ECCLS document 1986; 3(2).
- (21) Treskes M, Adriaansen HJ, van der Leur SJ, Idema RN, Pequeriaux N, Pronk C. Multicentre evaluation of the EBIO plus glucose analyser. Eur J Clin Chem Clin Biochem 1996; 34(9):777-784.
- (22) Burrin JM, Alberti KGMM. What is blood Glucose: can it be measured? Diabet Med 1990; 7:199-206.
- (23) Arens S, Moons V, Meuleman P, Struyf F, Zaman Z. Evaluation of Glucocard Memory 2 and Accutrend sensor blood glucose meters. Clin Chem Lab Med 1998; 36(1):47-52.
- (24) Weitgasser R, Straberger A, Schnoll F, Sailer S. Klinische Beurteilung der Blutzuckermeßgeräte Accutrend, Companion 2, Glucometer 3 and One Touch II im Vergleich mit der Referenzmethode Glukoseoxidase. Wien Klin Wochenschr 1994; 106(23):738-741.
- (25) Dillon AE, Menard MK, Rust P, Newman RB, VanDorsten JP. Glucometer analysis of one-hour glucose challenge samples. Am J Obstet Gynecol 1997; 177(5):1120-1123.
- (26) Grunert C, Wood WG. Comparison between the Bayer ELITE glucometer and the hexokinase method for glucose determination on the Eppendorf EPOS 5060 using capillary whole blood samples. Eur J Clin Chem Clin Biochem 1995; 33(3):153-156.
- (27) Girouard J, Forest JC, Massé J, Leroux M, Bradburn NC, Noblet TC, Ostrom Joynes J, Baum J. Multicenter evaluation of the Glucometer ELITE® XL meter, an instrument specifically designed for use with neonates. Diabetes Care 2000; 23:1149-1153.
- (28) Sorger M, Deperschmid M, Salvey E, Schlebusch H. Ascensia ELITE® XL – Evaluation des jüngsten Modells aus der bewährten Ascensia®-Familie. Ascensia(R) Diabetes Service - Bayer Health Care 2003;1-5.
- (29) Phillipou G, Seaborn CJ, Hooper J, Phillips PJ. Capillary blood glucose measurements in hospital inpatients using portable glucose meters. Aust N Z J Med 1993; 23(6):667-671.
- (30) Wiener K. An assessment of the effect of haematocrit on the HemoCue blood glucose analyzer. Ann Clin Biochem 1993; 30:90-93.
- (31) Schlebusch H, Niesen M, Sorger M, Paffenholz I, Fahnenstich H. Blood glucose determinations in newborns: four instruments compared. Pediatr Pathol Lab Med 1998; 18(1):41-48.
- (32) Lefevre G, Girardot-Dubois S, Chevallier G, Cohen C, Couderc R, Etienne J. Evaluation of the quality of blood glucose meters using the HemoCue B glucose system. Diabetes Metab 1999; 25(4):350-355.

- (33) Ickeringill M. The HemoCue B-Haemoglobin Analyser. Update in Anaesthesia 2001; 13:9-10.
- (34) Stork AD, Kemperman H, Erkelens DW, Veneman TF. Comparison of the accuracy of the HemoCue glucose analyzer with the Yellow Springs Instrument glucose oxidase analyzer, particularly in hypoglycemia. Eur J Endocrinol 2005; 153(2):275-281.
- (35) Palmer SM, Kaufman RA, Salamone SJ, Blake-Court, Bette W, Wahl HP, Furrer F. Cobas Integra: clinical laboratory instrument with continuous and random-access capabilities. Clin Chem 1995; 41(12 Pt 1):1751-1760.
- (36) Lippi G, Brentegani C, Mazzi C, Recchi C, Ruzzenente O, Guidi G. Evaluation of five new liquid stable applications on the Roche Cobas Integra. Eur J Clin Chem Clin Biochem 1997; 35(11):877-880.
- (37) Muser J, Bienvenu J, Blanckaert N, Brandslund I, Delattre J, Soffiati G, Swaminathan R, Maggini S, Mastall H. Inter-laboratory evaluation of the Cobas Integra 400 analytical system. Clin Chem Lab Med 2001; 39(6):539-559.
- (38) Redondo FL, Bermudez P, Cocco C, Colella F, Graziani MS, Fiehn W, Hierla T, Lemoel G, Belliard A, Manene D, Meziani M, Liebel M, McQueen MJ, Stockmann W. Evaluation of Cobas Integra 800 under simulated routine conditions in six laboratories. Clin Chem Lab Med 2003; 41(3):365-381.
- (39) Roche Diagnostics. Cobas Integra 400/700/800 - Glukose HK. Roche-Broschüre 2004.
- (40) Kunst A, Draeger B, Ziegenhorn J. UV-methods with hexokinase and glucose-6-phosphate dehydrogenase. In: Bergmeyer HU, editor. Methods of Enzymatic Analysis. Weinheim: Verlag Chemie, 1984: 163-172.
- (41) Thomas L. Blutglukose. In: Thomas L, editor. Labor und Diagnose. Frankfurt/Main: TH-Books Verlagsgesellschaft mbH, 1998.
- (42) World Health Organisation Study Group on Prevention of Diabetes mellitus. Technical Report Series 844. WHO document 1994.
- (43) Cederholm J, Wibell L. Insulin release and peripheral sensitivity an the oral glucose tolerance test. Diabetes Res Clin Pract 1990; 10:167-175.
- (44) Mari A, Pacini G, Murphy E, Ludvik B, Nolan JJ. A model-based method for assessing insulin sensitivity from the oral glucose tolerance test. Diabetes Care 2001; 24(3):539-548.
- (45) Perry RC, Shankar RR, Fineberg N, McGill J, Baron A, Adams PC. HbA1c measurement improves the detection of type 2 diabetes in high-risk individuals with nondiagnostic levels of fasting plasma glucose. Diabetes Care 2001; 24(3):465-471.
- (46) Isomaa B, Almgren P, Tuomi T, Forsen B, Lahti K, Nissen M, Taskinen MR, Groop L. Cardiovascular morbidity and mortality associated with the metabolic syndrome. Diabetes Care 2001; 24(4):683-689.
- (47) Alberti KG, Zimmet PZ. Definition, diagnosis and classification of diabetes mellitus and its complications. Part 1: Diagnosis and classification of diabetes mellitus. Provisional report of a WHO consultation. Diabet Med 1998; 15(7):539-553.
- (48) Truschnig-Wilders M, Boehm BO, März W. Laboratoriumsdiagnostik bei Diabetes mellitus. J Lab Med 2004; 28(4):334-345.
- (49) Kerner W, Brückel J, Böhm BO. Evidenzbasierte Diabetes-Leitlinien der Deutschen Diabetes Gesellschaft (DDG). Definition, Klassifikation und Diagnostik des Diabetes mellitus. Deutsche Diabetes-Gesellschaft 2004; Aktualisierung der 1. Auflage:1-11.

- (50) Passing H, Bablok W. A new biometrical procedure for testing the equality of measurements from two different analytical methods. Application of linear regression procedures for method comparison studies in clinical chemistry, Part I. *J Clin Chem Clin Biochem* 1983; 21(11):709-720.
- (51) Bablok W, Haeckel R, Meyers W, Wosniok W. Biometrical methods. In: Haeckel R, editor. *Evaluation Methods in Laboratory Medicine*. Weinheim, Germany: VCH, 1993: 203-241.
- (52) Barembruch BW, Stockmann W, Brauer P, Grager P, Michel R, Vonderschmitt D. CAEv- a program for computer-aided evaluation. *J Autom Chem* 1991; 13:167-179.
- (53) Silverman BW. *Density estimation for statistics and data analysis*. London: Chapman and Hall, 1990.
- (54) Jones MC, Marron JS, Sheather SJ. A brief survey of band width selection for density estimation. *J Amer Stat Ass* 1996; 91:401-407.
- (55) Brooks KE, Rawal N, Henderson AR. Laboratory assessment of three new monitors of blood glucose: Accu-Chek II, Glucometer II, and Glucoscan 2000. *Clin Chem* 1986; 32(12):2195-2200.
- (56) Chen ET, Nichols JH, Duh SH, Hortin G. Performance evaluation of blood glucose monitoring devices. *Diabetes Technol Ther* 2003; 5(5):749-768.
- (57) Fogh-Andersen N, D'Orazio P. Proposal for standardizing direct-reading biosensors for blood glucose. *Clin Chem* 1998; 44(3):655-659.
- (58) Kristensen GB, Christensen NG, Thue G, Sandberg S. Between-lot variation in external quality assessment of glucose: clinical importance and effect on participant performance evaluation. *Clin Chem* 2005; 51(9):1632-1636.
- (59) Skeie S, Thue G, Nerhus K, Sandberg S. Instruments for self-monitoring of blood glucose: comparisons of testing quality achieved by patients and a technician. *Clin Chem* 2002; 48(7):994-1003.
- (60) Petersen PH, Ricos C, Stockl D, Libeer JC, Baadenhuijsen H, Fraser C, Thienpont L. Proposed guidelines for the internal quality control of analytical results in the medical laboratory. *Eur J Clin Chem Clin Biochem* 1996; 34(12):983-999.
- (61) Petersen PH, Stöckl D, Westgard JO, Sandberg S, Linnet K, Thienpont L. Models for combining random and systematic errors. Assumptions and consequences for different models. *Clin Chem Lab Med* 2001; 39(7):589-595.
- (62) Price CP, Burrin JM, Natrass M. Extra-laboratory blood glucose measurement: a policy statement. *Diabet Med* 1988; 5(7):705-709.
- (63) Fraser CG, Petersen PH. Desirable standards for laboratory tests if they are to fulfill medical needs. *Clin Chem* 1993; 39(7):1447-1453.
- (64) Weiss SL, Cembrowski GS, Mazze RS. Patient and physician analytic goals for self-monitoring blood glucose instruments. *Am J Clin Pathol* 1994; 102(5):611-615.
- (65) Kost GJ, Vu HT, Lee JH, Bourgeois P, Kiechle FL, Martin C. Multicentre study of oxygen-insensitive handheld glucose point-of-care testing in the United States and Canada. *Crit Care Med* 1998; 26:581-590.
- (66) Ricos C, Juvany R, Simon M, Hernandez A, Alvarez V, Jimenez CV, Minchinela J, Perich C. Commutability and traceability: their repercussions on analytical bias and inaccuracy. *Clin Chim Acta* 1999; 280(1-2):135-145.

- (67) Skeie S, Thue G, Sandberg S. Patient-derived quality specifications for instruments used in self-monitoring of blood glucose. *Clin Chem* 2001; 47(1):67-73.
- (68) Sacks DB, Bruns DE, Goldstein DE, Maclaren NK, McDonald JM, Parrott M. Guidelines and recommendations for laboratory analysis in the diagnosis and management of diabetes mellitus. *Clin Chem* 2002; 48(3):436-472.
- (69) Bundesärztekammer. Richtlinien der Bundesärztekammer zur Qualitätssicherung in medizinischen Laboratorien. Bundesärztekammer Dokument 1993;38 Seiten.
- (70) Bundesärztekammer. Richtlinie der Bundesärztekammer zur Qualitätssicherung laboratoriumsmedizinischer Untersuchungen. *Deutsches Ärzteblatt* 2008; 105(7):A341-A355.
- (71) American Diabetes Association. Consensus statement on self-monitoring of blood glucose. *Diabetes Care* 1987; 10:93-99.
- (72) American Diabetes Association. Self-monitoring of blood glucose. *Diabetes Care* 1996; 19(Suppl. 1):62-66.
- (73) US Department of Health and Human Services. Clinical Laboratory Improvement Amendments of 1988: final rules and notice. 42 CFR Part 493. *Fed Regist* 1992; 57:7188-7288.
- (74) Sacks DB, Bernhardt P, Dunka LJ, Goldstein DE, Hortin GL, Mueller P, NCCLS. Point-of-Care Blood Glucose Testing in Acute and Chronic Care Facilities; Approved Guideline-Second Edition C30-A2. NCCLS document 2002; 22(17):28 Seiten.

8 ANHANG

8.1 TABELLEN

Tabelle 1: Maximal zulässige Fehlerrate für POCT-Glukometer, postuliert von Experten und Consensus-Gruppen

Individuelle Gruppen	Jahr	Totale Fehlerrate	Messbereich in mg/dl ^a
Cox <i>et al.</i> [4]	1985	≤ 20 %	
Clarke <i>et al.</i> [5]	1987	bezogen auf Zone A (error grid) : ≤ 20 % oder innerhalb des hypoglykämischen Bereiches, also ≤ 70 mg/dl	> 70 ≤ 70
Price <i>et al.</i> [62]	1988	≤ 10 %	
Koschinsky <i>et al.</i> [6]	1988	- 14 + 17 -35 % und +85 %	≤ 30 ≤ 350
Fraser <i>et al.</i> [63]	1993	≤ 1,9 %	
Weiss <i>et al.</i> [64]	1994	< + 15 % < + 20 %	> 70 ≤ 70
Kost <i>et al.</i> [65]	1998	< + 15 < + 25 %	< 100 ≥ 100
Ricos <i>et al.</i> [66]	1999	7,9%	
Skeie <i>et al.</i> [67]	2001		
Sacks <i>et al.</i> [68]	2002	≤ 7,9%	

Consensus-Gruppen	Jahr	Totale Fehlerrate	Messbereich in mg/dl
RiliBÄK '93 [69]	1993	< + 15 %	k.A.
RiliBÄK '02 [19]	2002	< + 15 % < 9 mg/dl	≥ 60 < 60
RiliBÄK '08 [70]	2008	< + 11 %	40 - 400
ADA'87 [7;71]	1987	≤ + 10 %	30 – 400
ADA'96 [7;72]	1996	< + 5 %	30 – 400
CLIA'88 [7;73]	1992	< + 10 % oder < + 6 mg/dl	k.A.
FDA und ISO-TC 212	1998	< + 20 % < + 20 mg/dl	≥ 100 < 100
NCCLS [74]	2002	< + 20 % ≤ 15 mg/dl	> 75 ≤ 75

^{a)} mmol/L-Angaben wurden durch den Faktor 18,02 in mg/dl umgerechnet

RiliBÄK = Richtlinie der Bundesärztekammer, ADA = American Diabetes Association, CLIA = Clinical Laboratory Improvement Amendments, FDA = Food and Drug Administration, ISO = International Organization for Standardization, NCCLS = National Committee for Clinical Laboratory Standards

Tabelle 2: "Hierarchy of Objectivity" (postuliert durch Fraser et Petersen [3])

1. Assessment of the effect of analytical performance on clinical decision making
Quality specifications in specific clinical situations
General quality specifications based on medical needs: e.g. based on biological variation
2. Professional recommendations from expert or institutional groups
3. Quality specifications laid down by regulation or by external quality assessment schemes (EQAS) organizers
4. Published data on the state of the art from proficiency testing schemes or published methodology

→ Zwei Stufen der Evaluation bei Methodenvergleichen

1. Stufe Definition der analytischen (Dis-)Übereinstimmung (konventionelles Konzept)
2. Stufe Schätzung der diagnostischen Relevanz (zukünftige Aufgabe/ Konzept)

Tabelle 8: Diagnostische Kriterien hinsichtlich des Glukosestoffwechsels entsprechend der Empfehlungen der WHO 1999 [42;47;48]; Glukosekonzentrationen in mg/dl (aus den 2004 überarbeiteten Evidenzbasierten Leitlinien Diabetes mellitus der Deutschen Diabetes Gesellschaft [49])

	Plasmaglukose, venös			Vollblutglukose, kapillär			Vollblutglukose, venös		
	Nüchtern		2h - oGTT	Nüchtern		2h - oGTT	Nüchtern		2h - oGTT
NGT Normale Glukose- toleranz	< 100	und	< 140	< 90	und	< 140	< 90	und	< 120
IFG Abnorme Nüchtern- glukose	≥ 100 < 126		---	≥ 90 < 110		---	≥ 90 < 110		---
IGT gestörte Glukose- toleranz	< 126	und	≥ 140 < 200	< 110	und	≥ 140 < 200	< 110	und	≥ 120 < 180
DM Diabetes mellitus	≥ 126	und/ oder	≥ 200	≥ 110	und/ oder	≥ 200	≥ 110	und/ oder	≥ 180

Nüchtern ist definiert durch eine Fastenperiode von wenigstens 8 Stunden.

Sofern keine ausgeprägte Hyperglykämie mit metabolischer Dekompensation vorliegt, ist die Diagnose durch eine Messung an einem anderen Tag zu bestätigen.

Venöses Plasma: bevorzugte Messung

oGTT (oraler Glukose-Toleranz-Test): 75 g Glukose in 300 ml Wasser über 3-5 Minuten

IFG: Impaired Fasting Glukose, IGT: Impaired Glukosetolerance

Tabelle 11: Kenngrößen diagnostischer Tests

Kenngrößen	auf Vierfeldertafel beruhend	als bedingte Wahrscheinlichkeit	ausformuliert
Prävalenz:		$P(K+)$	Wahrscheinlichkeit, mit der die Krankheit vorliegt
Sensitivität:	$A / A + C$	$P(T+ K+)$	Kranke aus den Kranken erkennen (Testergebnis richtig positiv)
Spezifität:	$D / B + D$	$P(T- K-)$	Nicht-Kranke aus den Nicht-Kranken erkennen (Testergebnis richtig negativ)
Effizienz:	$A + D / A + B + C + D$		Richtig-Getestete aus allen erkennen
In-Effizienz:	$B + C / A + B + C + D$		Falsch-Getestete aus allen erkennen
PW(pos):	$A / A + B$	$P(K+ T+)$	Kranke unter den Krank-Getesteten erkennen (Richtig-positiv-Getestete unter den Positiv-Getesteten)
PW(neg):	$D / C + D$	$P(K- T-)$	Nicht-Kranke unter den Nicht-krank-Getesteten erkennen (Richtig-negativ-Getestete unter den Negativ-Getesteten)
LR(pos):	$(A / A + C) / [1 - (D / B + D)]$		richtig positiv (von allen Erkrankten) / falsch positiv (von allen Gesunden)
LR(neg):	$[1 - (A / A + C)] / (D / B + D)$		falsch negativ (von allen Erkrankten) / richtig negativ (von allen Gesunden)
DR(pos)*:	$B / A + B + C + D$		Falsch-positiv-Getestete aus allen erkennen
DR(neg)*:	$C / A + B + C + D$		Falsch-negativ-Getestete aus allen erkennen
DR(ges)*~:	$B + C / A + B + C + D$		Falsch-Getestete aus allen erkennen

 DR = Diskordanzrate; ges = gesamt; K = Krankheit; LR = Likelihood Ratio; neg = negativ; P = Wahrscheinlichkeit; pos = positiv; PW = prädiktiver Wert; T = Testergebnis

*) unter Berücksichtigung der Wahrscheinlichkeit, dass eine Diskordanz aufgrund von Streuung der Messwerte auftritt, und der Wahrscheinlichkeit, dass ein Wert in einem repräsentativen Populationskollektivs überhaupt vorkommen kann.

~) DR(ges) entspricht nicht der In-Effizienz, auch wenn sich die abstrahierten Formeln entsprechen, weil bei Berechnung der DR(ges) Wahrscheinlichkeitsüberlegungen einfließen

Annahmen auf Vierfeldertafel beruhend

- A = richtig positiv
- B = falsch positiv
- C = falsch negativ
- D = richtig negativ

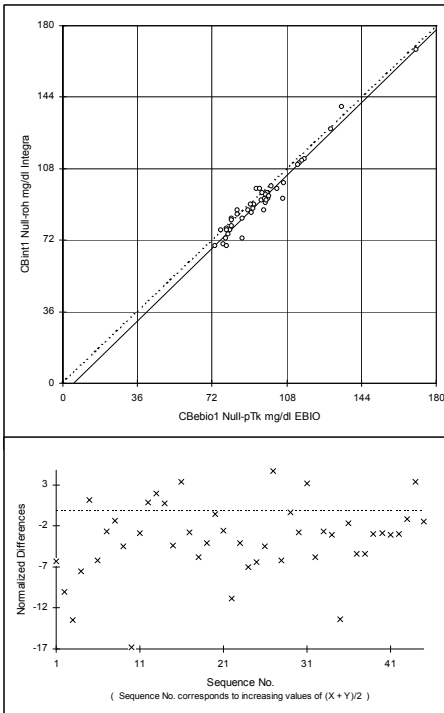
als bedingte Wahrscheinlichkeiten angegeben

- K+ = Krankheit vorhanden
- K- = Krankheit nicht vorhanden
- T+ = positives Testergebnis
- T- = negatives Testergebnis

8.2 ABBILDUNGEN

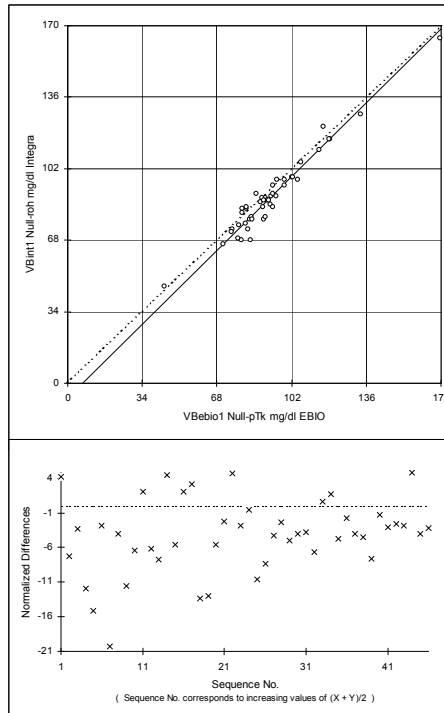
Abbildungen 12 a – f: Unkorrigierte Vergleiche von Integra vs. Ebio

12 a: Nüchtern-Kapillarblut



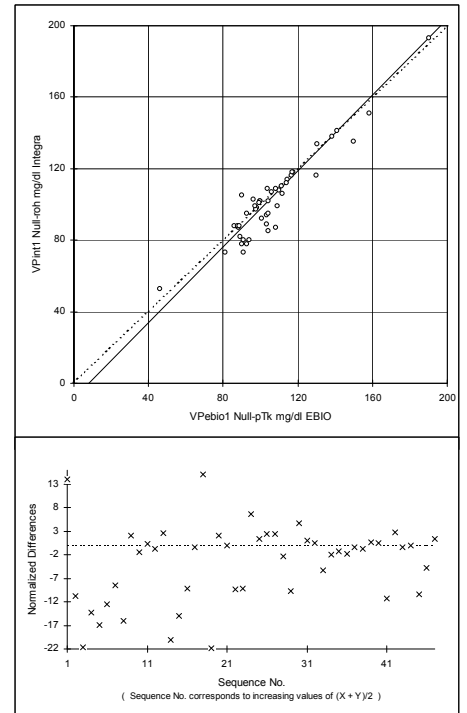
P/B Regression
 $Y = 1,019 * X - 5,363$
 $md(95) = 5,504$
 $N = 45, r = 0,977$

12 b: Nüchtern-Venenblut



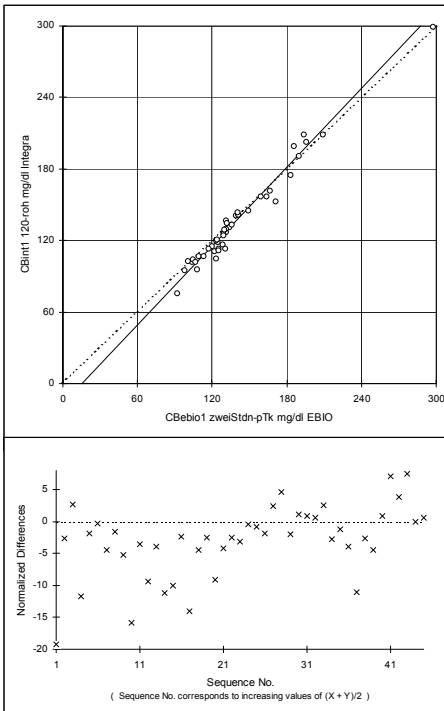
P/B Regression
 $Y = 1,032 * X - 7,170$
 $md(95) = 5,850$
 $N = 46, r = 0,974$

12 c: Nüchtern-Venenplasma



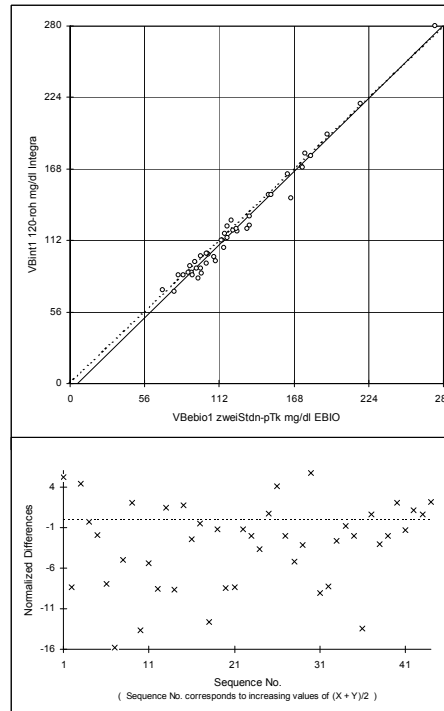
P/B Regression
 $Y = 1,061 * X - 8,628$
 $md(95) = 11,025$
 $N = 47, r = 0,945$

12 d: 2h-Postprandial-Kapillarblut



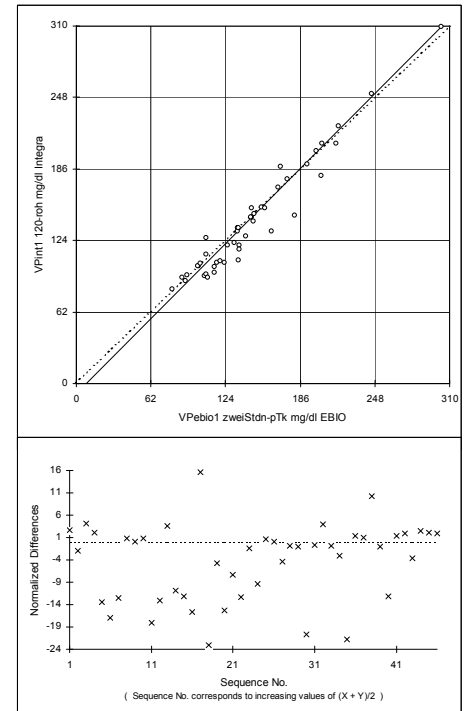
P/B Regression
 $Y = 1,102 * X - 16,571$
 $md(95) = 8,651$
 $N = 45, r = 0,986$

12 e: 2h-Postprandial-Venenblut



P/B Regression
 $Y = 1,024 * X - 6,022$
 $md(95) = 6,934$
 $N = 44, r = 0,991$

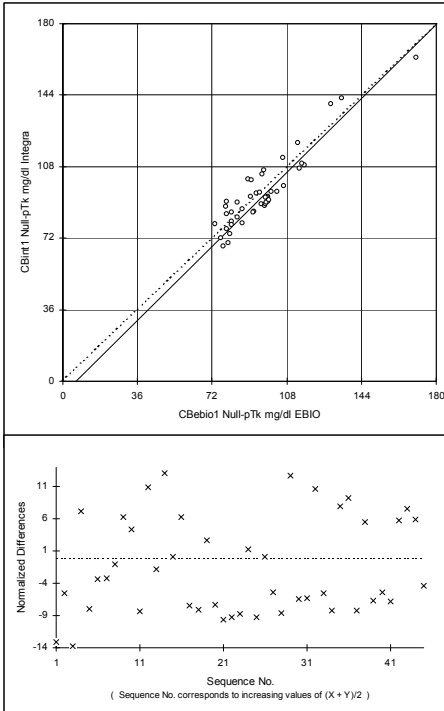
12 f: 2h-Postprandial-Venenplasma



P/B Regression
 $Y = 1,047 * X - 8,302$
 $md(95) = 17,485$
 $N = 46, r = 0,971$

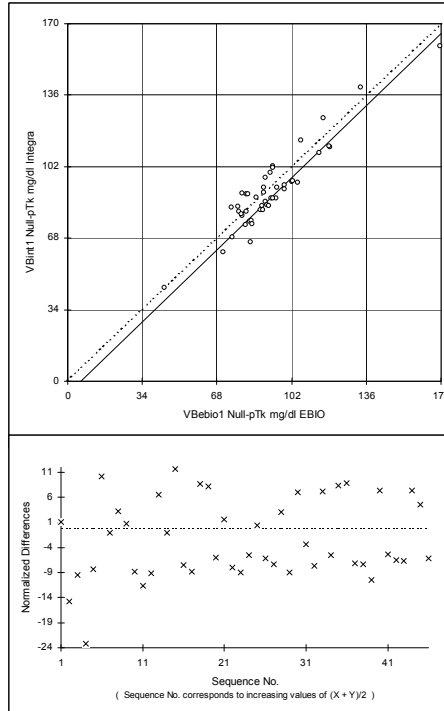
Abbildungen 13 a – f: Tages-korrigierte Vergleiche von Integra vs. Ebio

13 a: Nüchtern-Kapillarblut



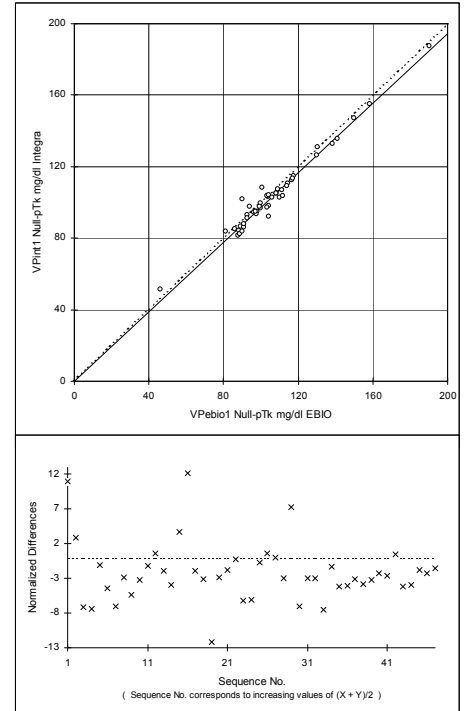
P/B Regression
 $Y = 1,034 * X - 6,616$
 $md(95) = 9,330$
 $N = 45, r = 0,926$

13 b: Nüchtern-Venenblut



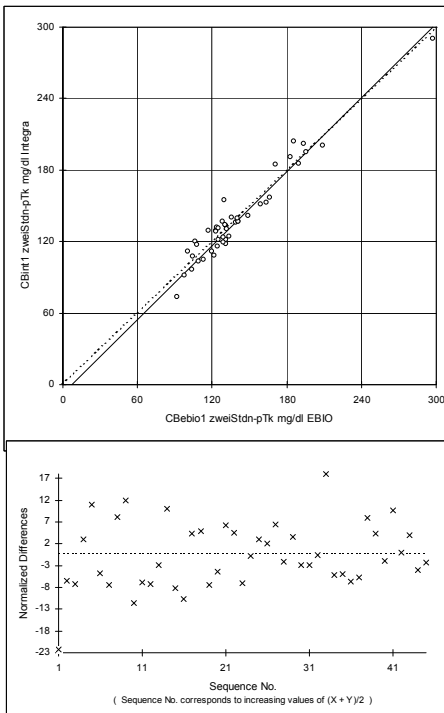
P/B Regression
 $Y = 1,008 * X - 5,987$
 $md(95) = 9,622$
 $N = 46, r = 0,934$

13 c: Nüchtern-Venenplasma



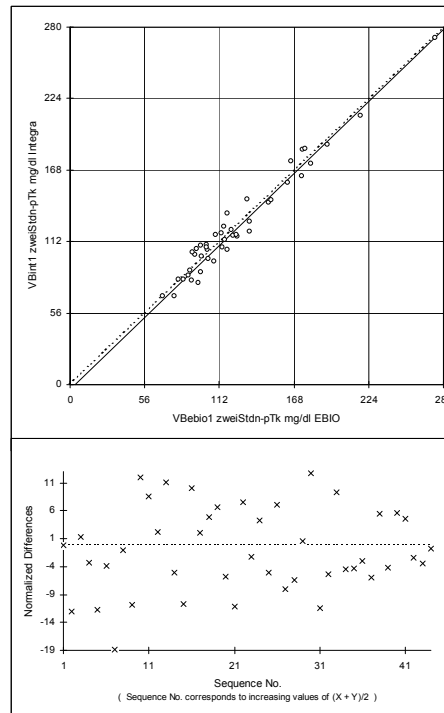
P/B Regression
 $Y = 0,973 * X - 0,084$
 $md(95) = 5,617$
 $N = 47, r = 0,985$

13 d: 2h-Postprandial-Kapillarblut



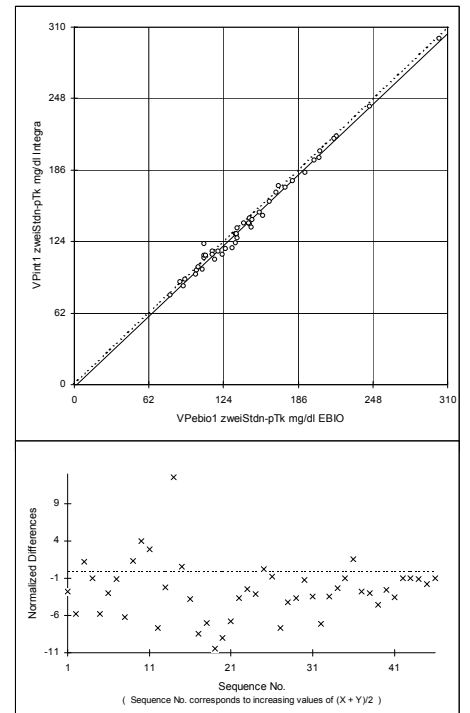
P/B Regression
 $Y = 1,035 * X - 7,825$
 $md(95) = 11,750$
 $N = 45, r = 0,968$

13 e: 2h-Postprandial-Venenblut



P/B Regression
 $Y = 1,005 * X - 3,492$
 $md(95) = 10,137$
 $N = 44, r = 0,979$

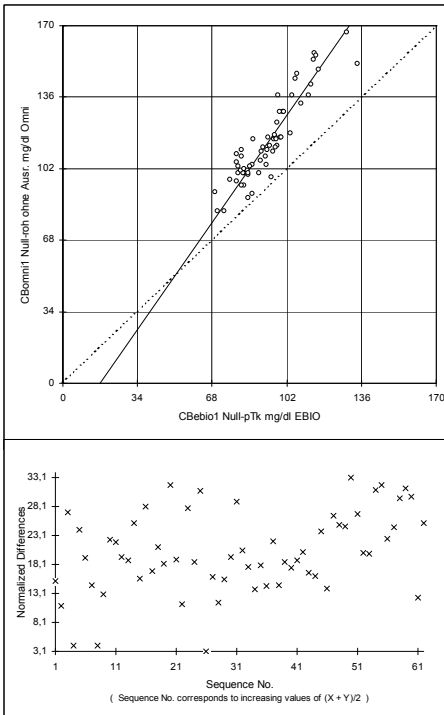
13 f: 2h-Postprandial-Venenplasma



P/B Regression
 $Y = 0,988 * X - 2,120$
 $md(95) = 5,567$
 $N = 46, r = 0,994$

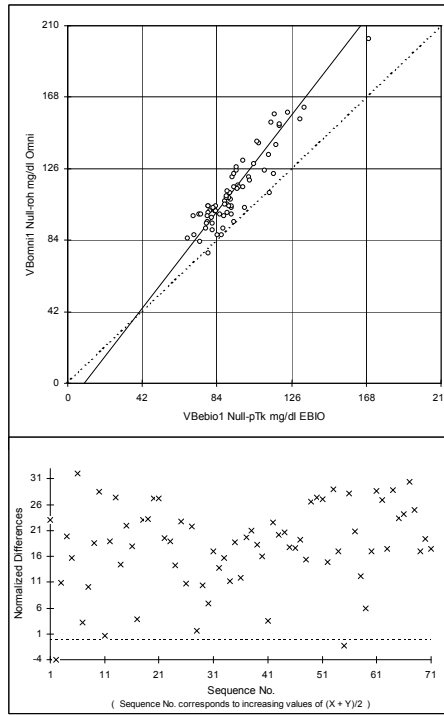
Abbildungen 14 a – f: Unkorrigierte Vergleiche von Omni vs. Ebio

14 a: Nüchtern-Kapillarblut



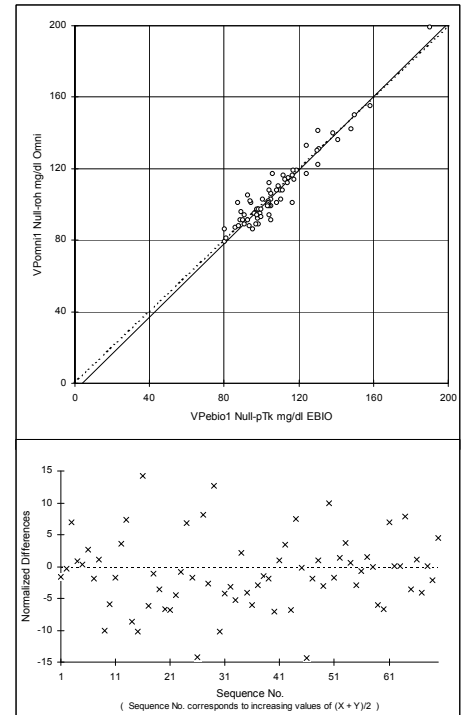
P/B Regression
 $Y = 1,497 * X - 25,155$
 $md(95) = 8,273$
 $N = 62, r = 0,915$

14 b: Nüchtern-Venenblut



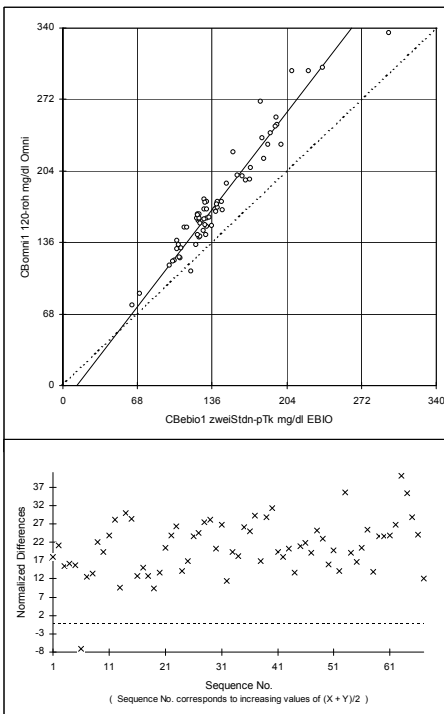
P/B Regression
 $Y = 1,349 * X - 12,619$
 $md(95) = 10,912$
 $N = 71, r = 0,920$

14 c: Nüchtern-Venenplasma



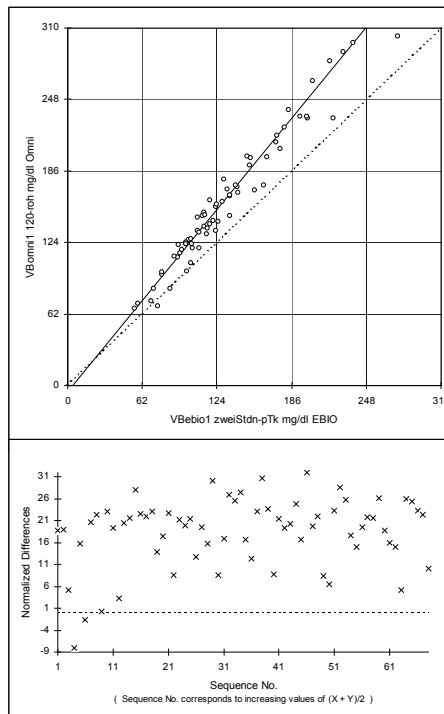
P/B Regression
 $Y = 1,030 * X - 4,337$
 $md(95) = 8,703$
 $N = 70, r = 0,955$

14 d: 2h-Postprandial-Kapillarblut



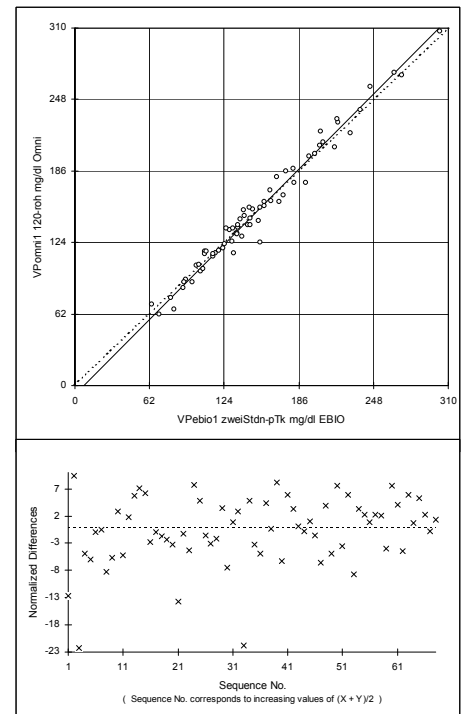
P/B Regression
 $Y = 1,360 * X - 17,626$
 $md(95) = 18,130$
 $N = 67, r = 0,964$

14 e: 2h-Postprandial-Venenblut



P/B Regression
 $Y = 1,271 * X - 4,880$
 $md(95) = 13,932$
 $N = 68, r = 0,977$

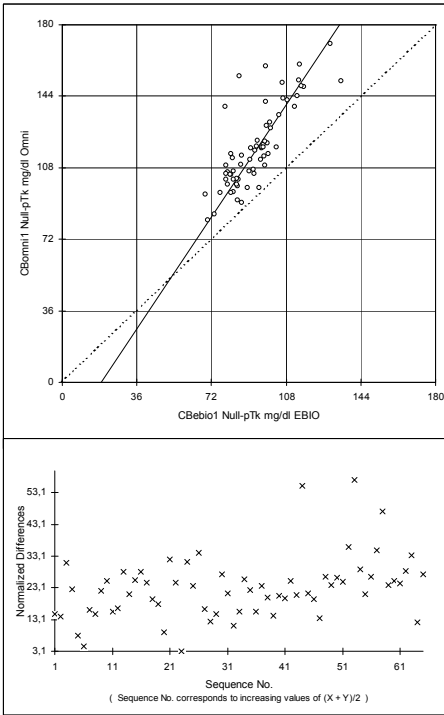
14 f: 2h-Postprandial-Venenplasma



P/B Regression
 $Y = 1,050 * X - 7,793$
 $md(95) = 9,335$
 $N = 68, r = 0,988$

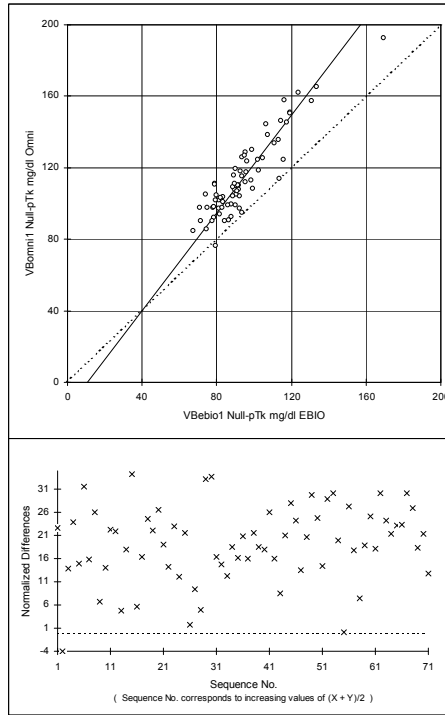
Abbildungen 15 a – f: Tages-korrigierte Vergleiche von Omni vs. Ebio

15 a: Nüchtern-Kapillarblut



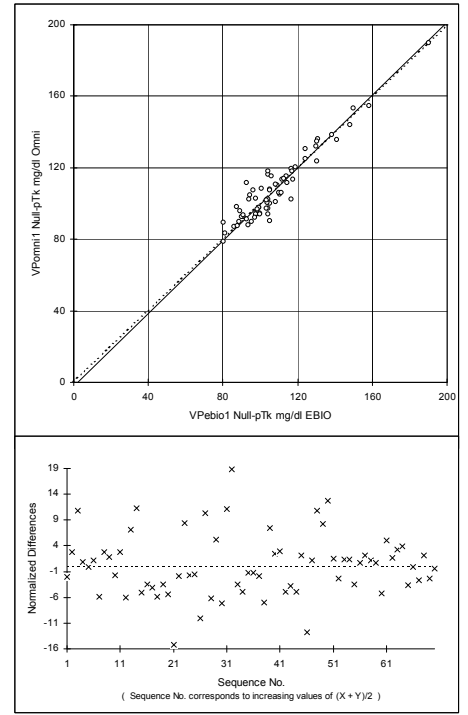
P/B Regression
 $Y = 1,562 * X - 29,026$
 $md(95) = 13,592$
 $N = 65, r = 0,808$

15 b: Nüchtern-Venenblut



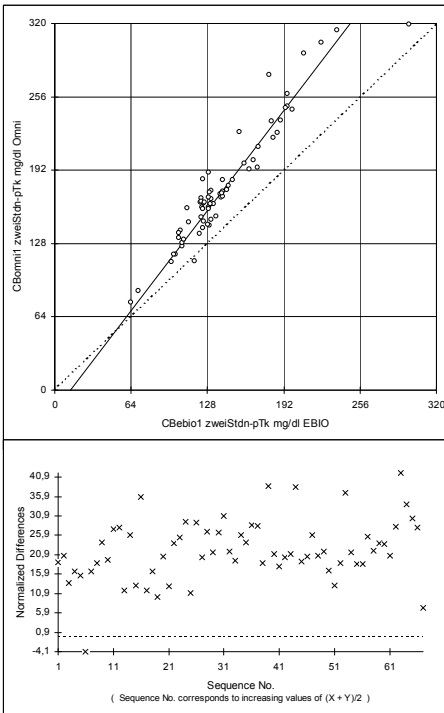
P/B Regression
 $Y = 1,364 * X - 14,311$
 $md(95) = 10,706$
 $N = 71, r = 0,911$

15 c: Nüchtern-Venenplasma



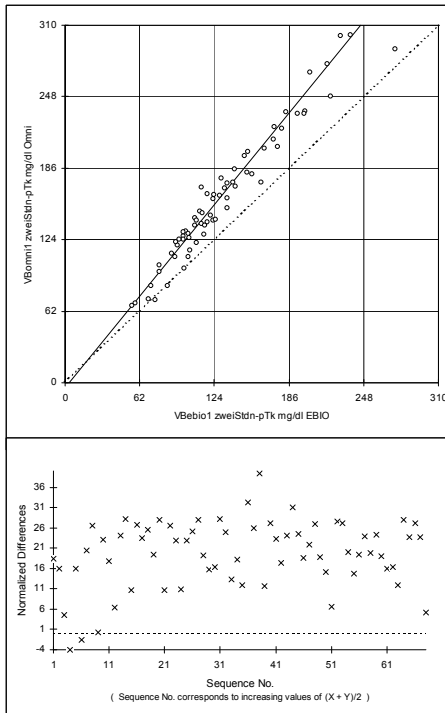
P/B Regression
 $Y = 1,018 * X - 2,119$
 $md(95) = 9,139$
 $N = 70, r = 0,947$

15 d: 2h-Postprandial-Kapillarblut



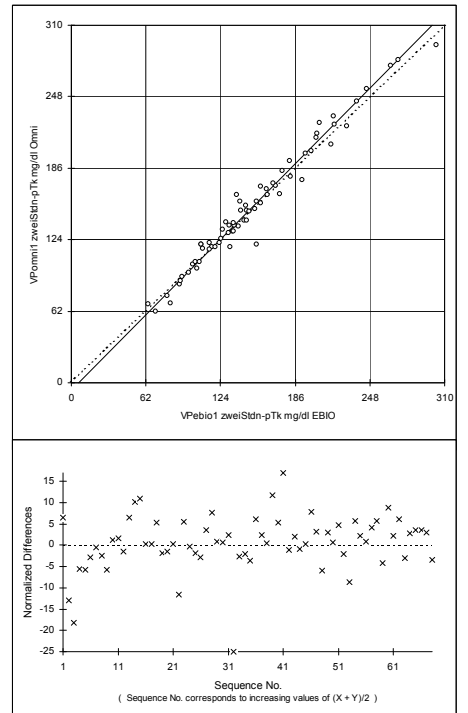
P/B Regression
 $Y = 1,364 * X - 17,681$
 $md(95) = 18,778$
 $N = 67, r = 0,954$

15 e: 2h-Postprandial-Venenblut



P/B Regression
 $Y = 1,280 * X - 4,480$
 $md(95) = 15,805$
 $N = 68, r = 0,974$

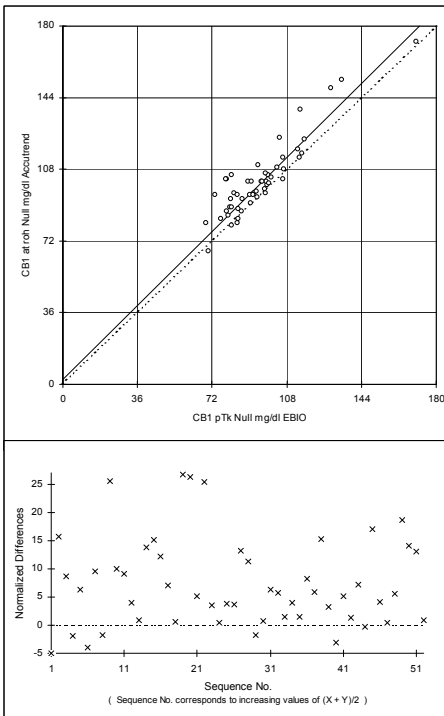
15 f: 2h-Postprandial-Venenplasma



P/B Regression
 $Y = 1,058 * X - 6,864$
 $md(95) = 12,459$
 $N = 68, r = 0,985$

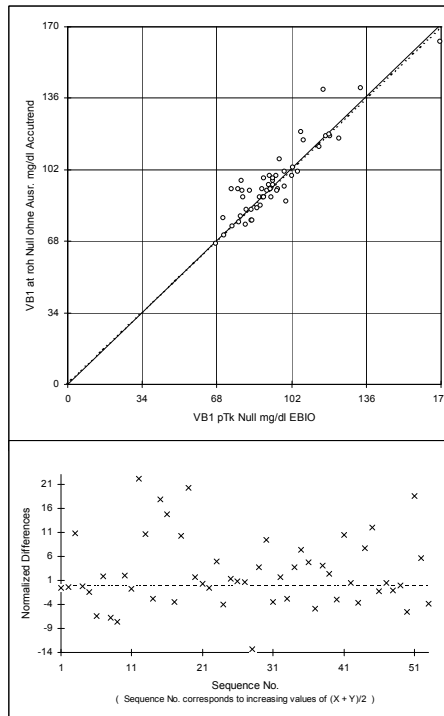
Abbildungen 16 a – f: Unkorrigierte Vergleiche von Accutrend vs. Ebio

16 a: Nüchtern-Kapillarblut



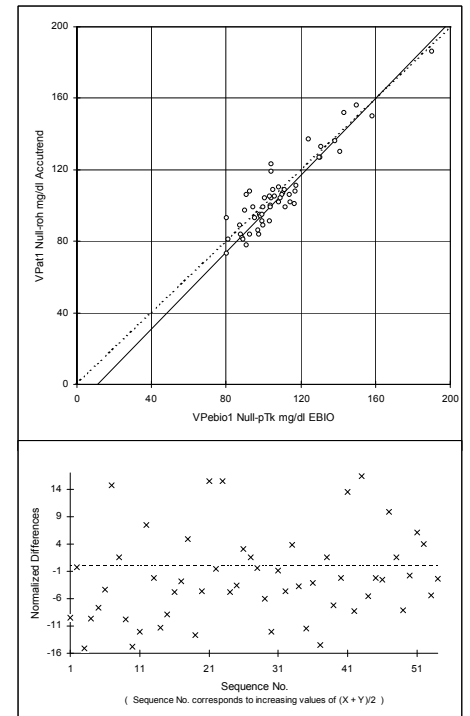
P/B Regression
 $Y = 1,032 * X + 2,565$
 $md(95) = 12,435$
 $N = 52, r = 0,913$

16 b: Nüchtern-Venenblut



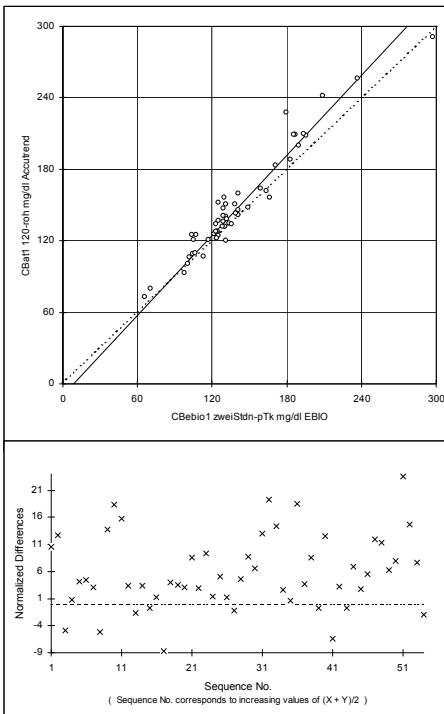
P/B Regression
 $Y = 1,008 * X - 0,142$
 $md(95) = 11,319$
 $N = 53, r = 0,921$

16 c: Nüchtern-Venenplasma



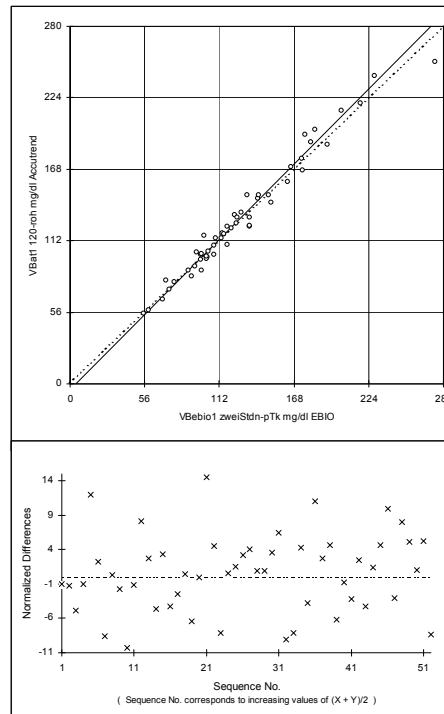
P/B Regression
 $Y = 1,075 * X - 12,033$
 $md(95) = 13,530$
 $N = 54, r = 0,926$

16 d: 2h-Postprandial-Kapillarblut



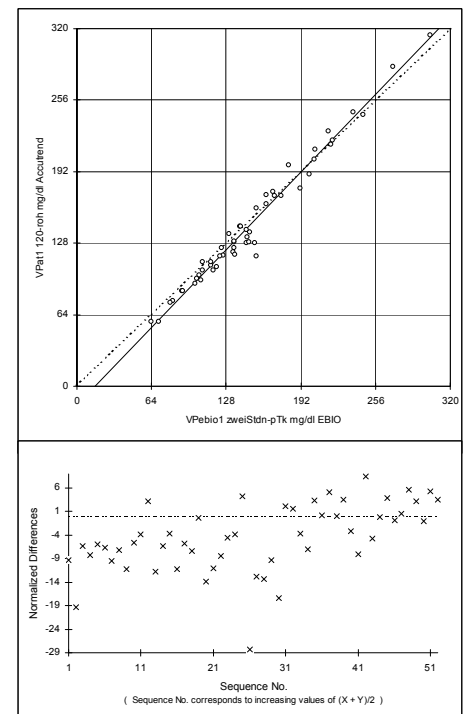
P/B Regression
 $Y = 1,119 * X - 9,548$
 $md(95) = 13,953$
 $N = 54, r = 0,968$

16 e: 2h-Postprandial-Venenblut



P/B Regression
 $Y = 1,051 * X - 4,699$
 $md(95) = 9,673$
 $N = 52, r = 0,985$

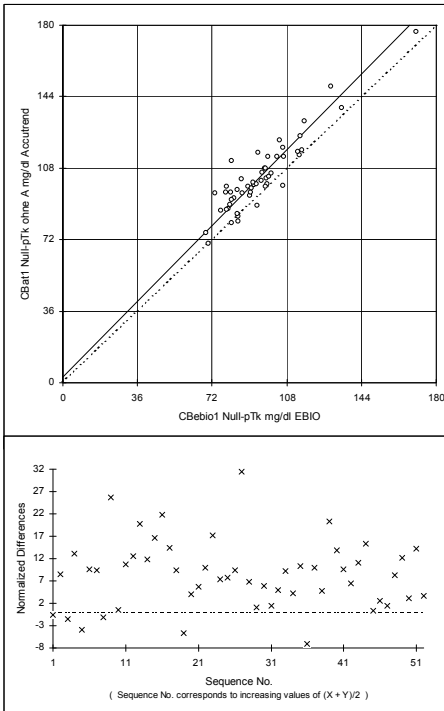
16 f: 2h-Postprandial-Venenplasma



P/B Regression
 $Y = 1,084 * X - 16,599$
 $md(95) = 10,828$
 $N = 52, r = 0,987$

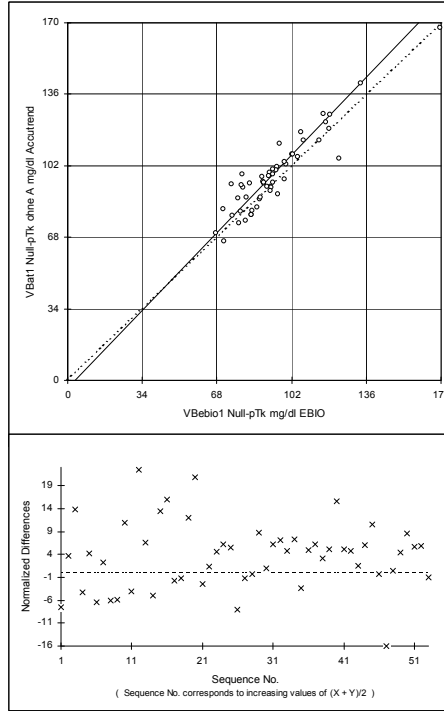
Abbildungen 17 a – f: Tages-korrigierte Vergleiche von Accutrend vs. Ebio

17 a: Nüchtern-Kapillarblut



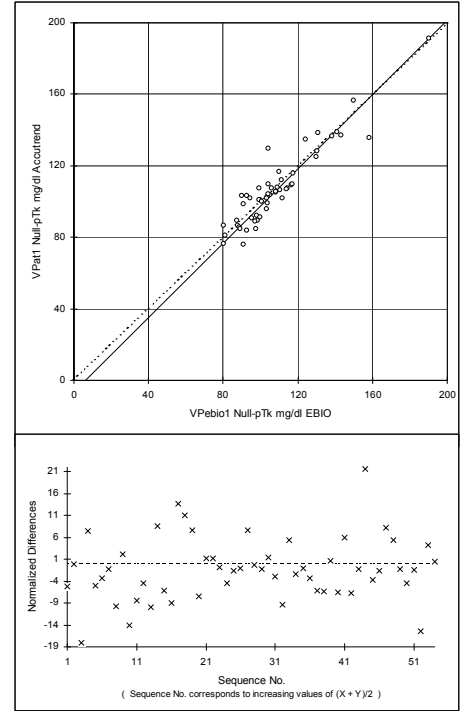
P/B Regression
 $Y = 1,059 * X + 3,087$
 $md(95) = 9,358$
 $N = 52, r = 0,919$

17 b: Nüchtern-Venenblut



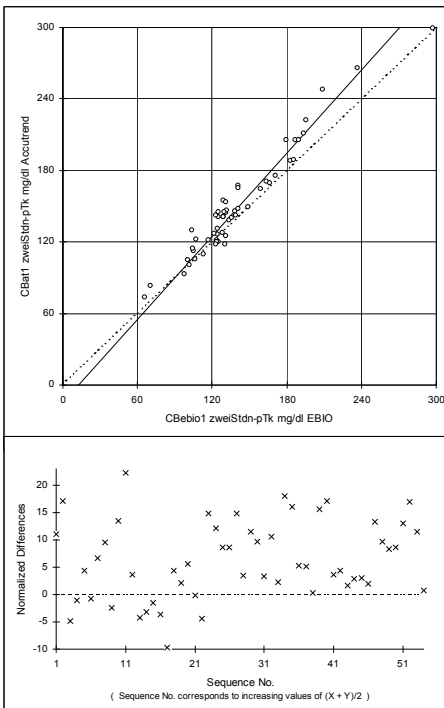
P/B Regression
 $Y = 1,088 * X - 3,677$
 $md(95) = 9,558$
 $N = 53, r = 0,927$

17 c: Nüchtern-Venenplasma



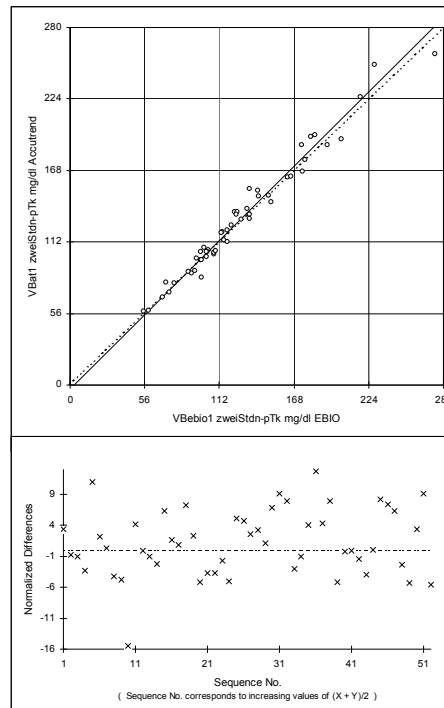
P/B Regression
 $Y = 1,038 * X - 6,370$
 $md(95) = 10,349$
 $N = 54, r = 0,933$

17 d: 2h-Postprandial-Kapillarblut



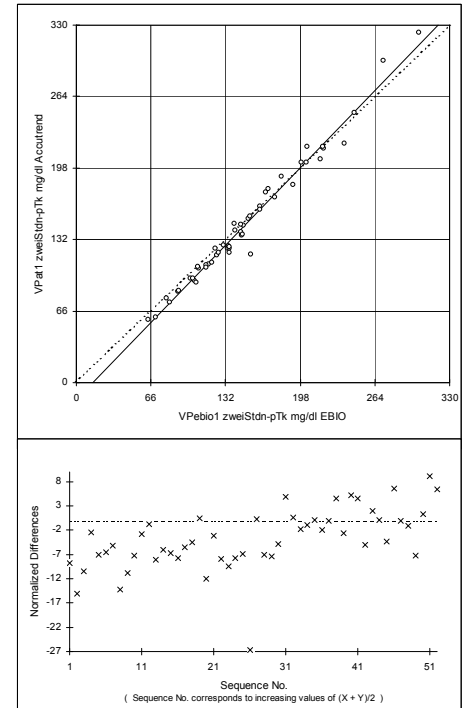
P/B Regression
 $Y = 1,163 * X - 14,478$
 $md(95) = 12,472$
 $N = 54, r = 0,972$

17 e: 2h-Postprandial-Venenblut



P/B Regression
 $Y = 1,037 * X - 2,971$
 $md(95) = 10,908$
 $N = 52, r = 0,986$

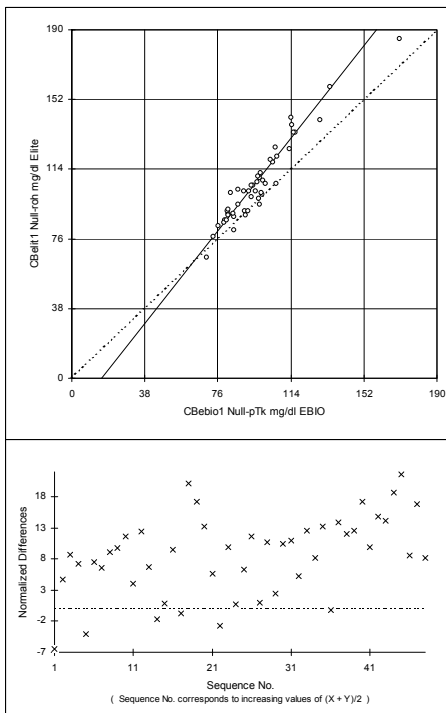
17 f: 2h-Postprandial-Venenplasma



P/B Regression
 $Y = 1,080 * X - 15,589$
 $md(95) = 11,154$
 $N = 52, r = 0,989$

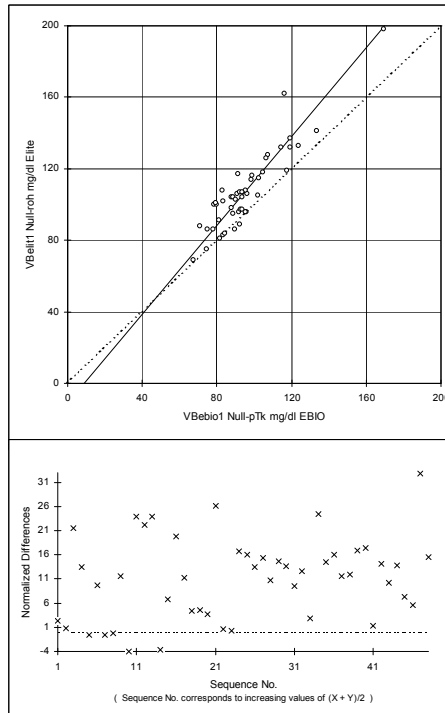
Abbildungen 18 a – f: Unkorrigierte Vergleiche von Elite vs. Ebio

18 a: Nüchtern-Kapillarblut



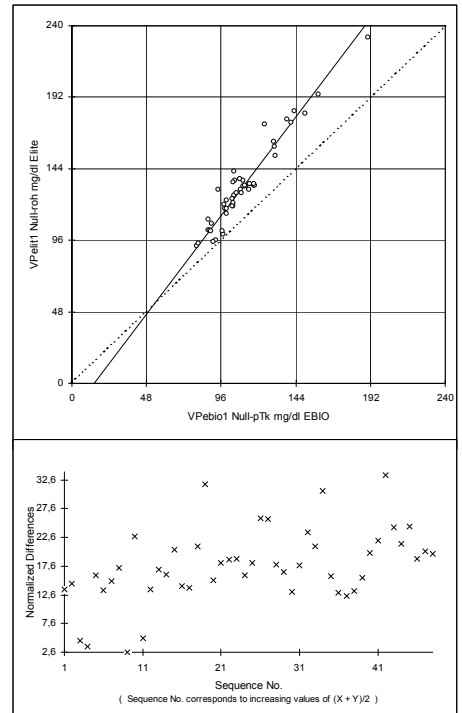
P/B Regression
 $Y = 1,330 * X - 20,844$
 $md(95) = 7,853$
 $N = 48, r = 0,957$

18 b: Nüchtern-Venenblut



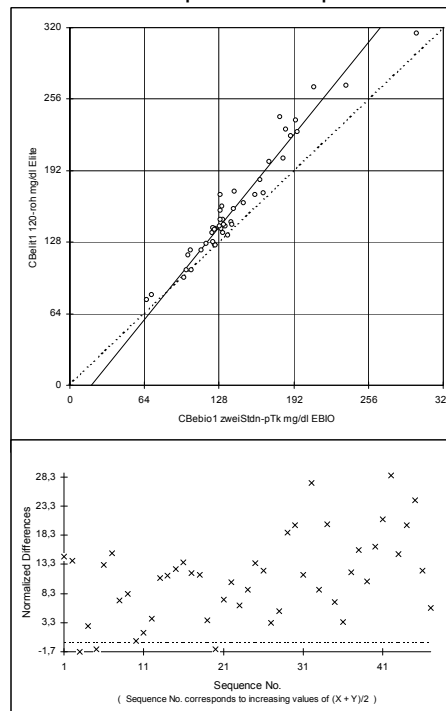
P/B Regression
 $Y = 1,240 * X - 10,836$
 $md(95) = 9,545$
 $N = 48, r = 0,915$

18 c: Nüchtern-Venenplasma



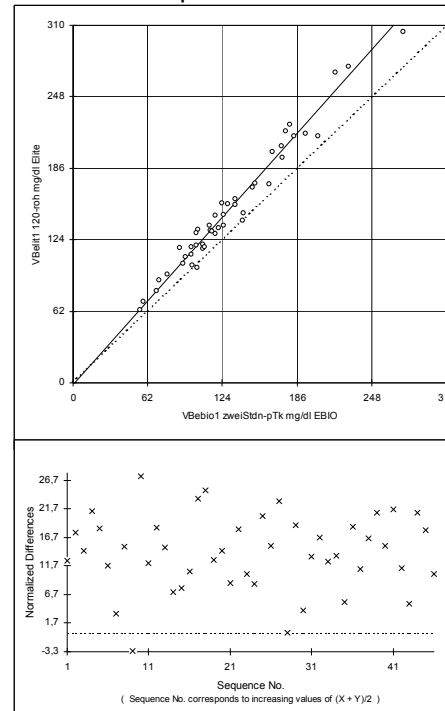
P/B Regression
 $Y = 1,380 * X - 19,891$
 $md(95) = 9,550$
 $N = 48, r = 0,963$

18 d: 2h-Postprandial-Kapillarblut



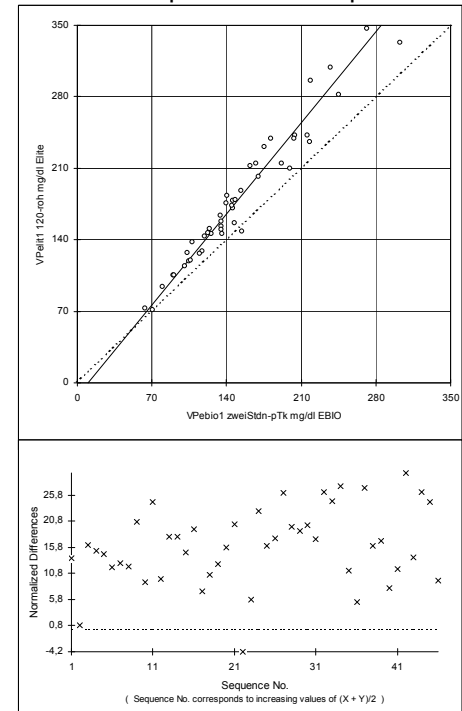
P/B Regression
 $Y = 1,289 * X - 23,345$
 $md(95) = 14,704$
 $N = 47, r = 0,970$

18 e: 2h-Postprandial-Venenblut



P/B Regression
 $Y = 1,174 * X - 1,889$
 $md(95) = 11,926$
 $N = 46, r = 0,983$

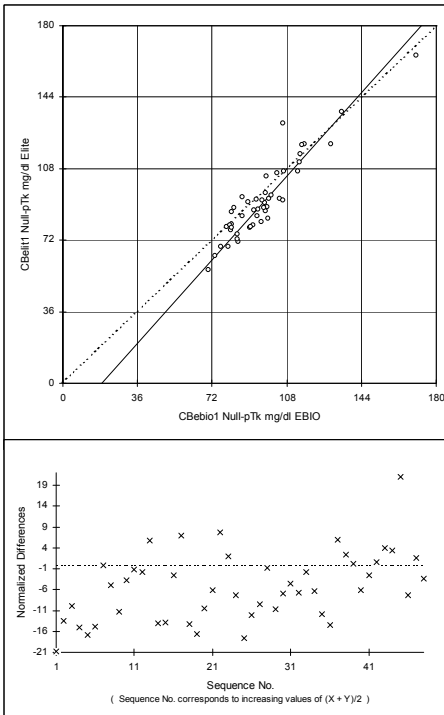
18 f: 2h-Postprandial-Venenplasma



P/B Regression
 $Y = 1,275 * X - 13,179$
 $md(95) = 18,780$
 $N = 46, r = 0,974$

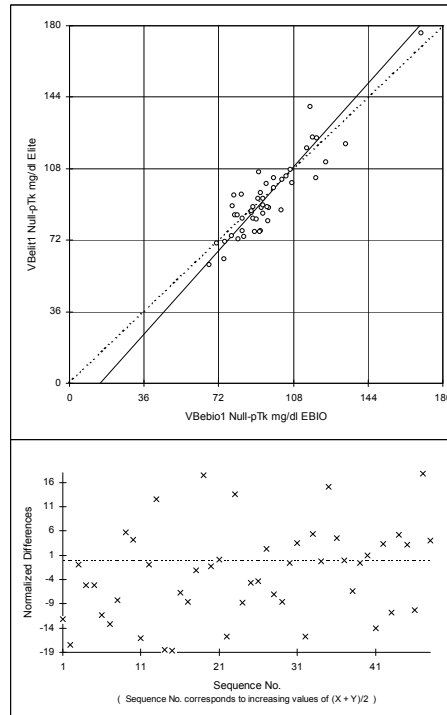
Abbildungen 19 a – f: Tages-korrigierte Vergleiche von Elite vs. Ebio

19 a: Nüchtern-Kapillarblut



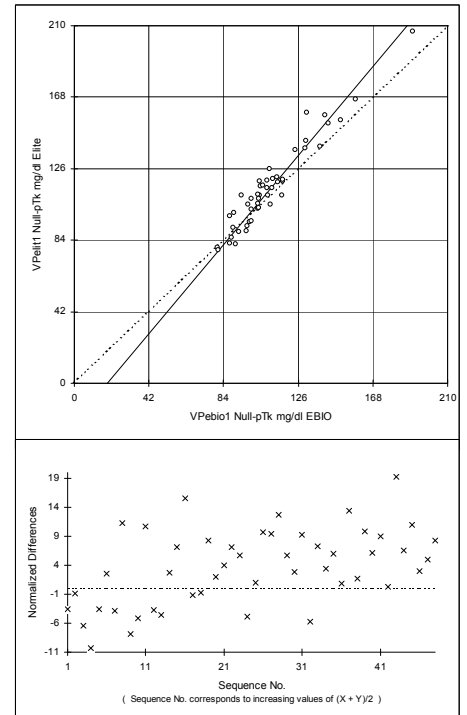
P/B Regression
 $Y = 1,167 * X - 21,838$
 $md(95) = 8,780$
 $N = 48, r = 0,927$

19 b: Nüchtern-Venenblut



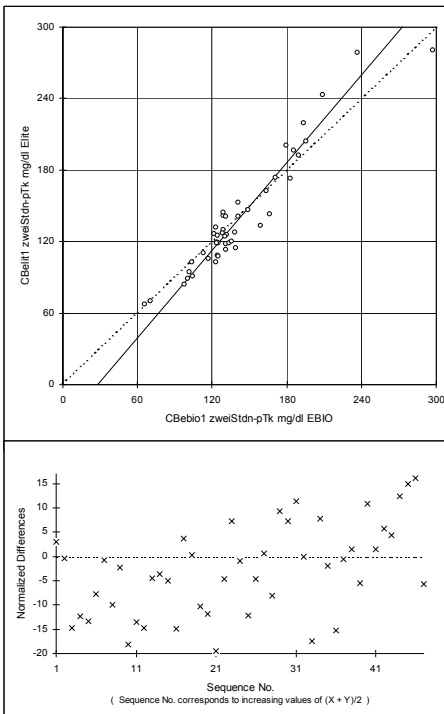
P/B Regression
 $Y = 1,168 * X - 17,154$
 $md(95) = 11,973$
 $N = 48, r = 0,902$

19 c: Nüchtern-Venenplasma



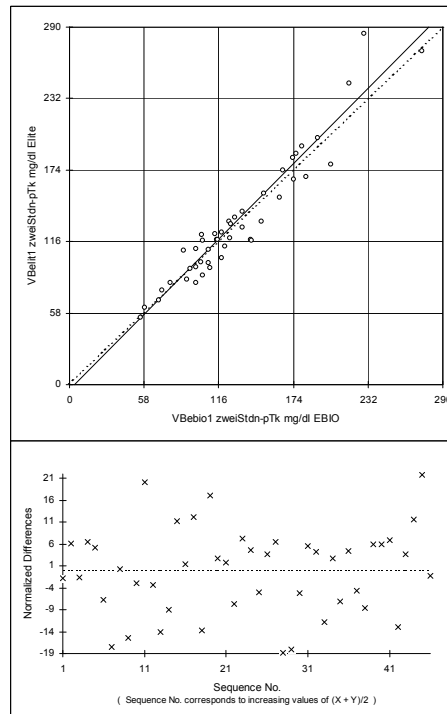
P/B Regression
 $Y = 1,244 * X - 22,899$
 $md(95) = 7,526$
 $N = 48, r = 0,962$

19 d: 2h-Postprandial-Kapillarblut



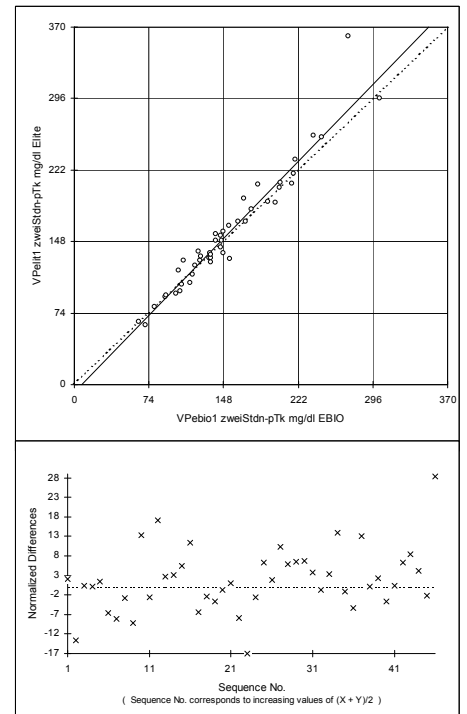
P/B Regression
 $Y = 1,228 * X - 34,522$
 $md(95) = 15,501$
 $N = 47, r = 0,958$

19 e: 2h-Postprandial-Venenblut



P/B Regression
 $Y = 1,056 * X - 4,379$
 $md(95) = 18,820$
 $N = 46, r = 0,959$

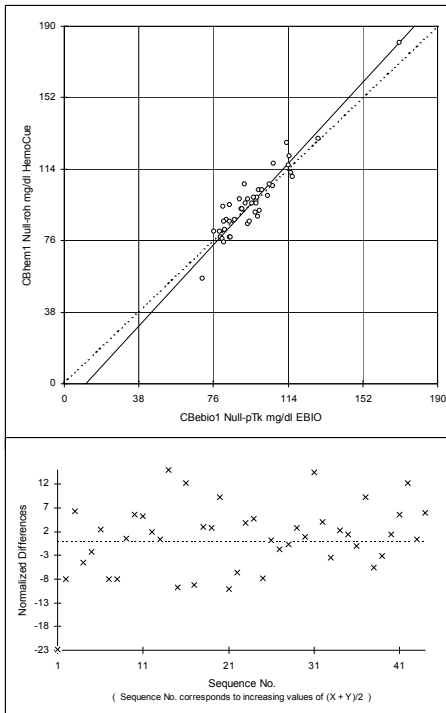
19 f: 2h-Postprandial-Venenplasma



P/B Regression
 $Y = 1,077 * X - 8,313$
 $md(95) = 14,191$
 $N = 46, r = 0,967$

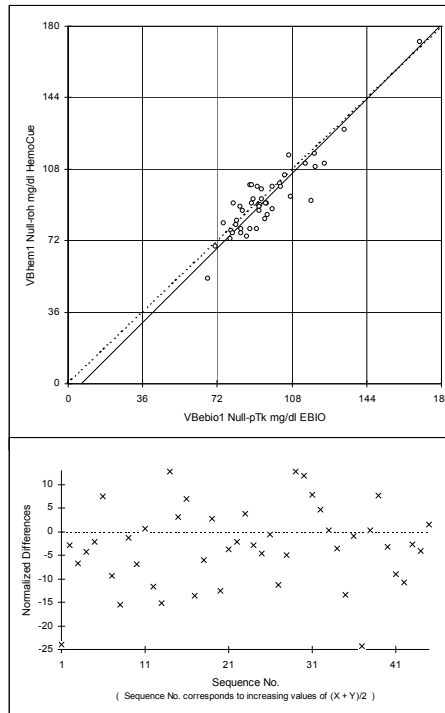
Abbildungen 20 a – f: Unkorrigierte Vergleiche von HemoCue vs. Ebio

20 a: Nüchtern-Kapillarblut



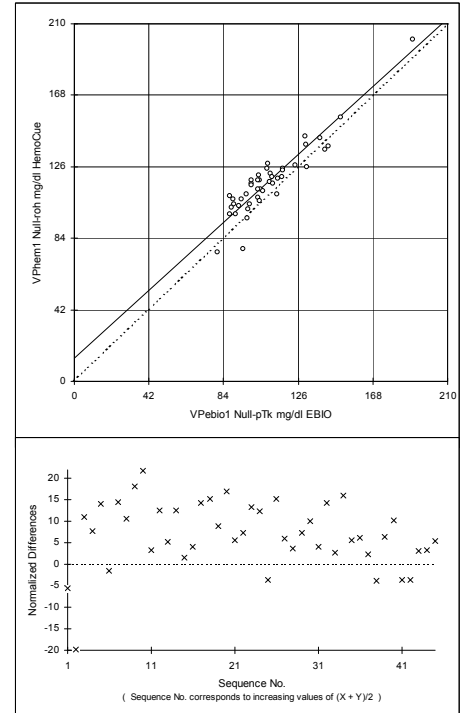
P/B Regression
 $Y = 1,138 * X - 12,686$
 $md(95) = 7,869$
 $N = 44, r = 0,942$

20 b: Nüchtern-Venenblut



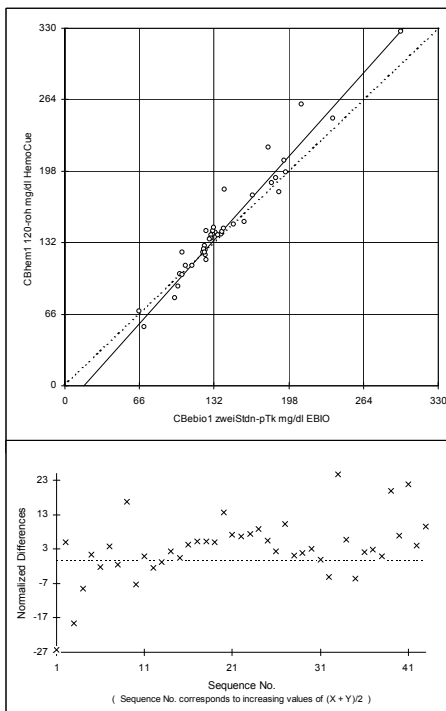
P/B Regression
 $Y = 1,043 * X - 6,792$
 $md(95) = 9,858$
 $N = 45, r = 0,910$

20 c: Nüchtern-Venenplasma



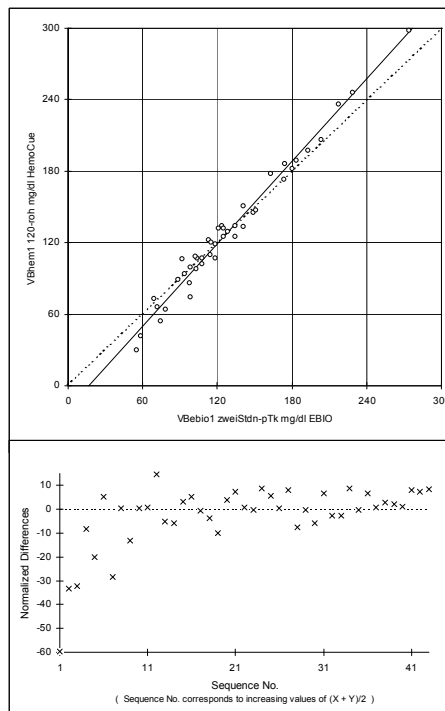
P/B Regression
 $Y = 0,950 * X + 13,491$
 $md(95) = 8,791$
 $N = 45, r = 0,923$

20 d: 2h-Postprandial-Kapillarblut



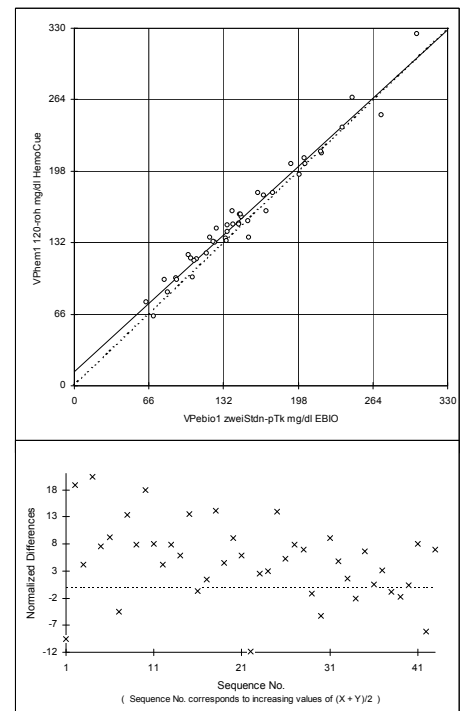
P/B Regression
 $Y = 1,168 * X - 19,372$
 $md(95) = 16,895$
 $N = 43, r = 0,971$

20 e: 2h-Postprandial-Venenblut



P/B Regression
 $Y = 1,153 * X - 18,963$
 $md(95) = 9,264$
 $N = 43, r = 0,988$

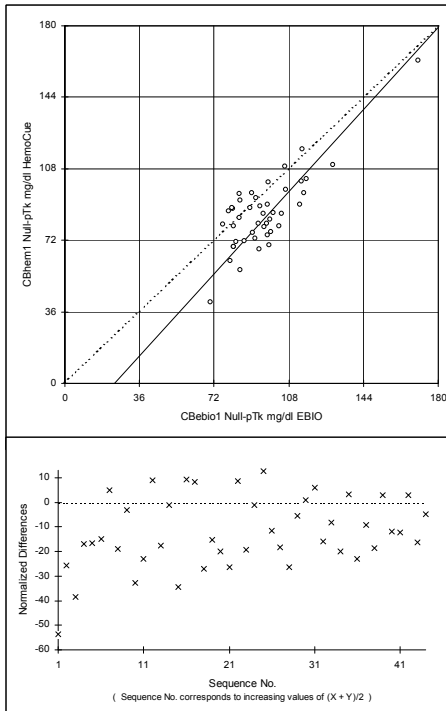
20 f: 2h-Postprandial-Venenplasma



P/B Regression
 $Y = 0,957 * X + 12,807$
 $md(95) = 14,519$
 $N = 43, r = 0,983$

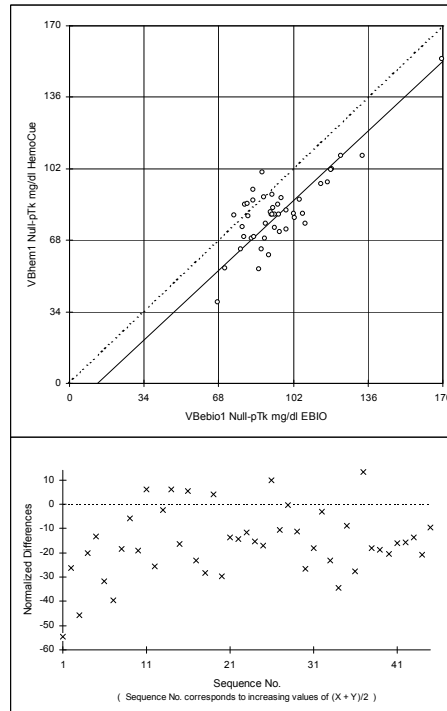
Abbildungen 21 a – f: Tages-korrigierte Vergleiche von HemoCue vs. Ebio

21 a: Nüchtern-Kapillarblut



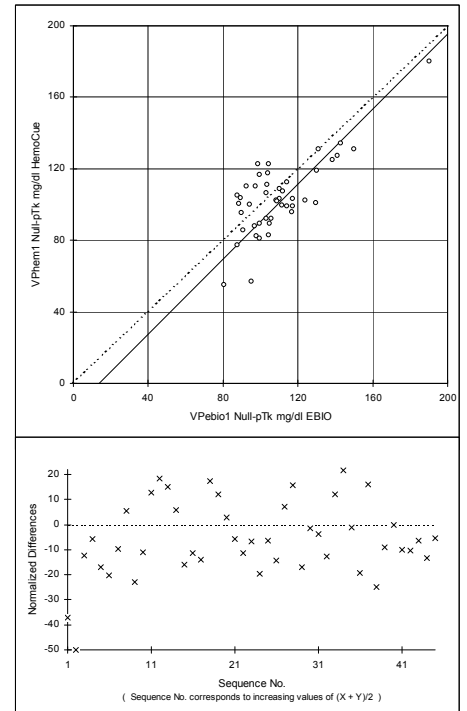
P/B Regression
 $Y = 1,149 * X - 27,558$
 $md(95) = 15,045$
 $N = 44, r = 0,797$

21 b: Nüchtern-Venenblut



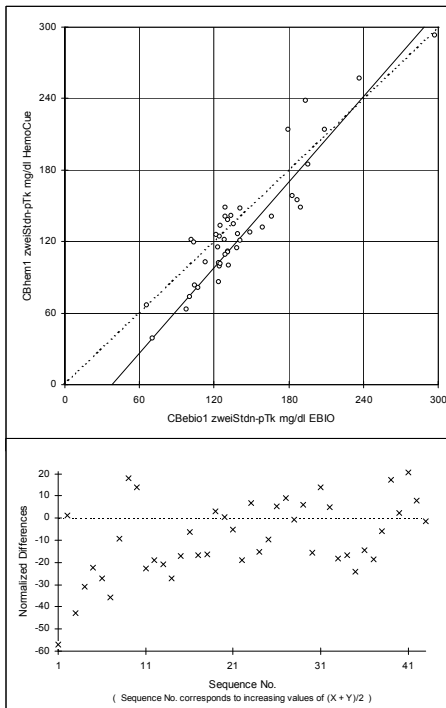
P/B Regression
 $Y = 0,975 * X - 12,517$
 $md(95) = 13,897$
 $N = 45, r = 0,813$

21 c: Nüchtern-Venenplasma



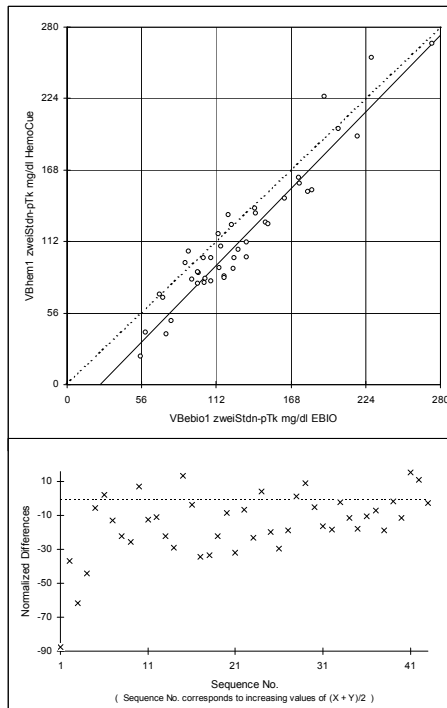
P/B Regression
 $Y = 1,048 * X - 14,428$
 $md(95) = 19,131$
 $N = 45, r = 0,758$

21 d: 2h-Postprandial-Kapillarblut



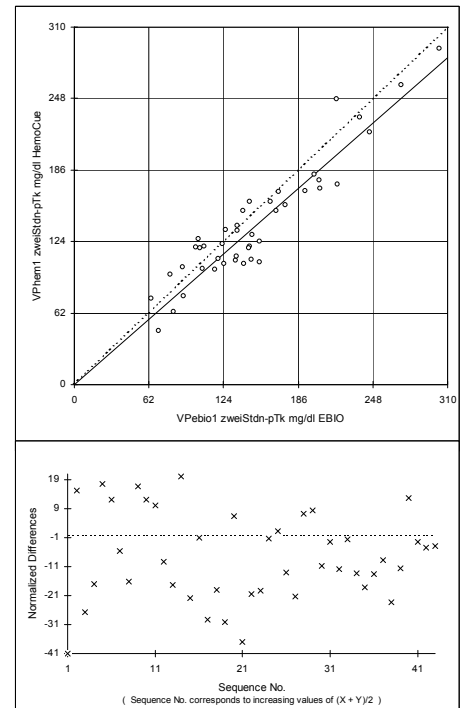
P/B Regression
 $Y = 1,196 * X - 45,030$
 $md(95) = 26,958$
 $N = 43, r = 0,919$

21 e: 2h-Postprandial-Venenblut



P/B Regression
 $Y = 1,072 * X - 26,453$
 $md(95) = 21,355$
 $N = 43, r = 0,952$

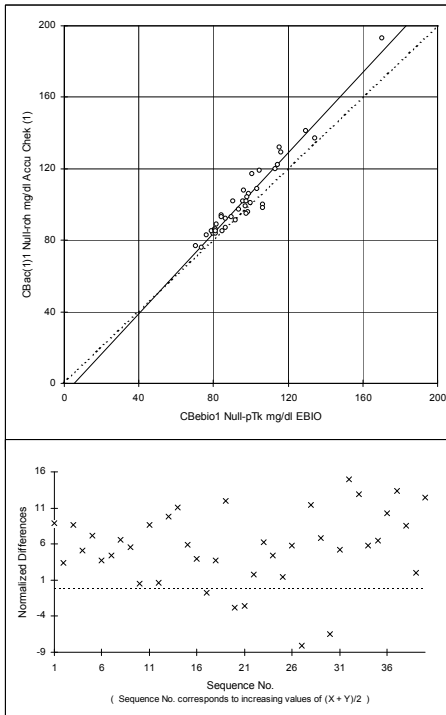
21 f: 2h-Postprandial-Venenplasma



P/B Regression
 $Y = 0,914 * X - 0,288$
 $md(95) = 21,903$
 $N = 43, r = 0,935$

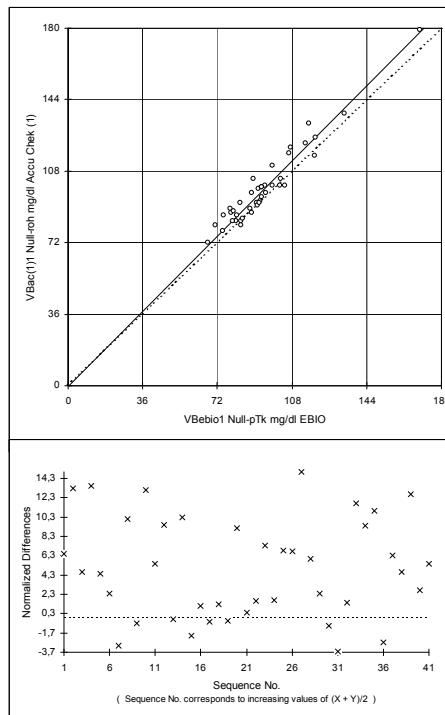
Abbildungen 22 a – f: Unkorrigierte Vergleiche von Accu-Chek1 vs. Ebio

22 a: Nüchtern-Kapillarblut



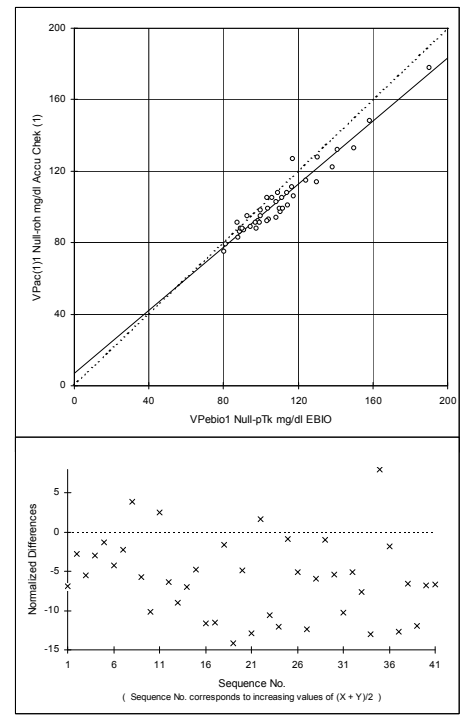
P/B Regression
 $Y = 1,125 * X - 6,047$
 $md(95) = 7,888$
 $N = 40, r = 0,966$

22 b: Nüchtern-Venenblut



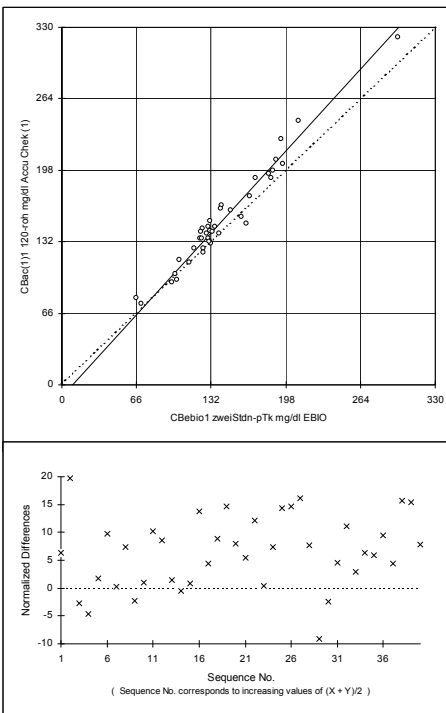
P/B Regression
 $Y = 1,052 * X - 0,319$
 $md(95) = 6,194$
 $N = 41, r = 0,966$

22 c: Nüchtern-Venenplasma



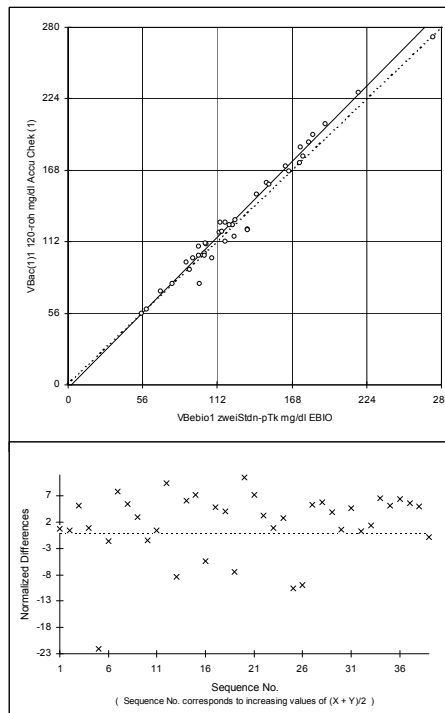
P/B Regression
 $Y = 0,881 * X + 7,096$
 $md(95) = 5,599$
 $N = 41, r = 0,966$

22 d: 2h-Postprandial-Kapillarblut



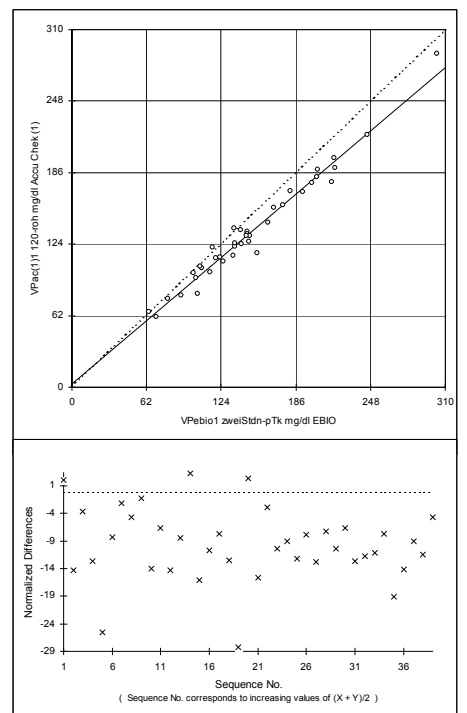
P/B Regression
 $Y = 1,145 * X - 10,707$
 $md(95) = 10,491$
 $N = 40, r = 0,979$

22 e: 2h-Postprandial-Venenblut



P/B Regression
 $Y = 1,057 * X - 2,724$
 $md(95) = 12,422$
 $N = 39, r = 0,988$

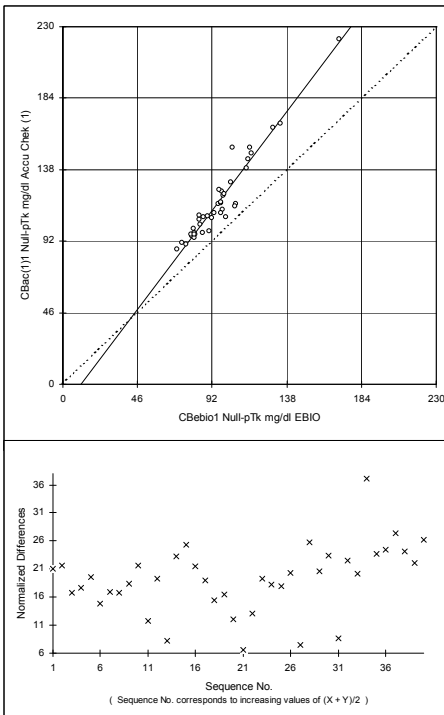
22 f: 2h-Postprandial-Venenplasma



P/B Regression
 $Y = 0,883 * X + 3,149$
 $md(95) = 12,718$
 $N = 39, r = 0,984$

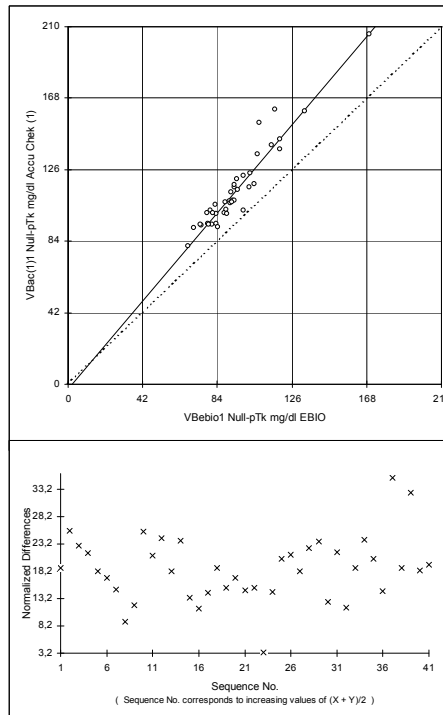
Abbildungen 23 a – f: Tages-korrigierte Vergleiche von Accu-Chek1 vs. Ebio

23 a: Nüchtern-Kapillarblut



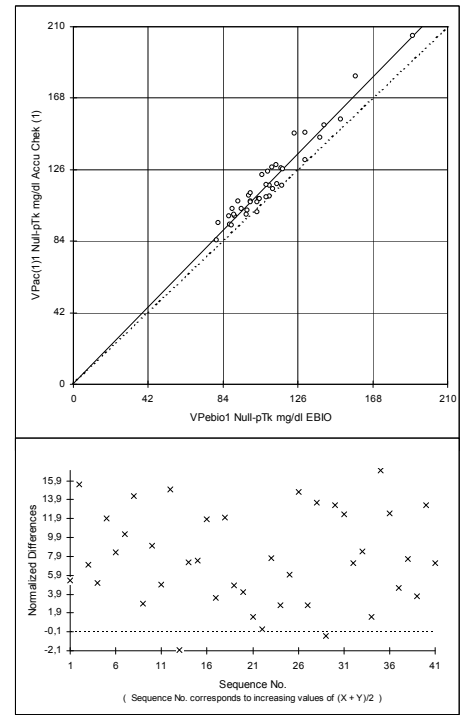
P/B Regression
 $Y = 1,382 * X - 15,181$
 $md(95) = 9,824$
 $N = 40, r = 0,962$

23 b: Nüchtern-Venenblut



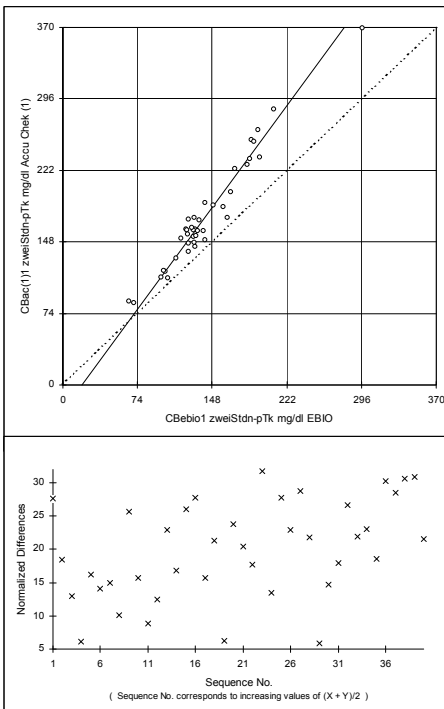
P/B Regression
 $Y = 1,231 * X - 2,656$
 $md(95) = 10,623$
 $N = 41, r = 0,954$

23 c: Nüchtern-Venenplasma



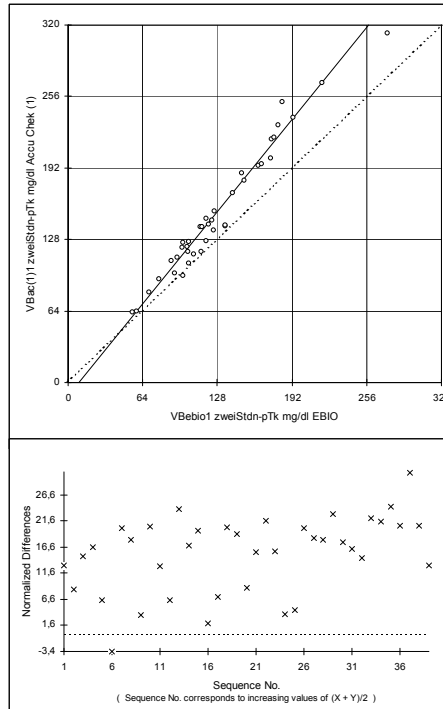
P/B Regression
 $Y = 1,073 * X + 0,402$
 $md(95) = 6,805$
 $N = 41, r = 0,970$

23 d: 2h-Postprandial-Kapillarblut



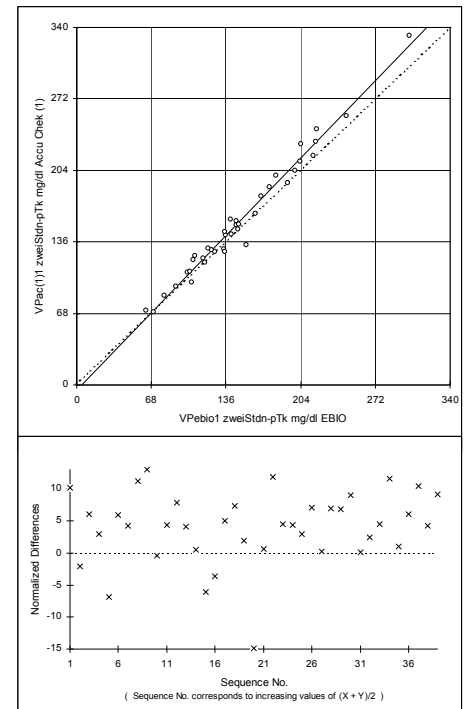
P/B Regression
 $Y = 1,426 * X - 27,451$
 $md(95) = 14,534$
 $N = 40, r = 0,975$

23 e: 2h-Postprandial-Venenblut



P/B Regression
 $Y = 1,290 * X - 11,767$
 $md(95) = 14,635$
 $N = 39, r = 0,982$

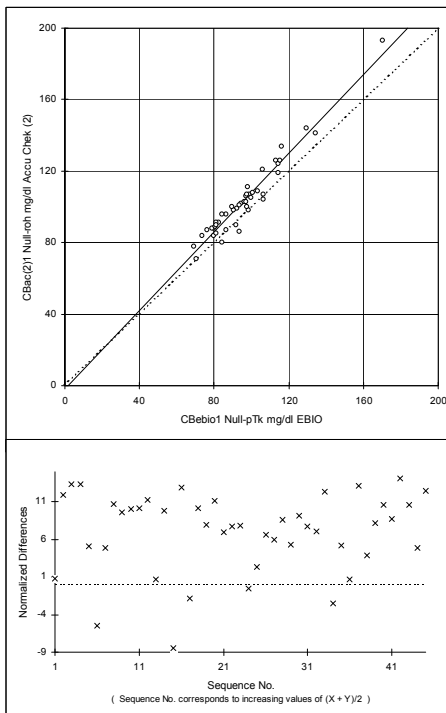
23 f: 2h-Postprandial-Venenplasma



P/B Regression
 $Y = 1,080 * X - 4,516$
 $md(95) = 9,357$
 $N = 39, r = 0,989$

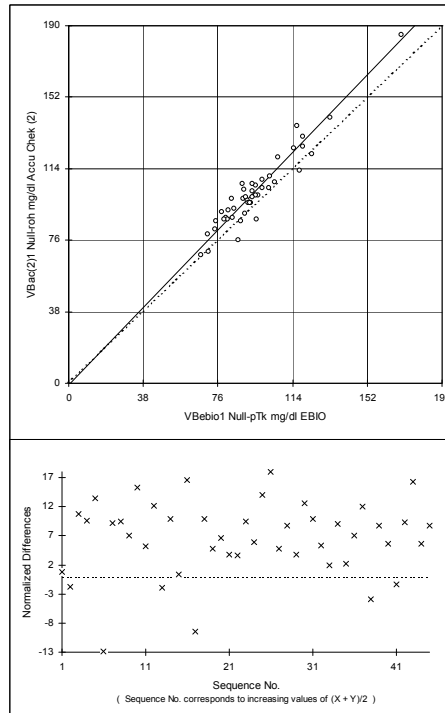
Abbildungen 24 a – f: Unkorrigierte Vergleiche von Accu-Chek2 vs. Ebio

24 a: Nüchtern-Kapillarblut



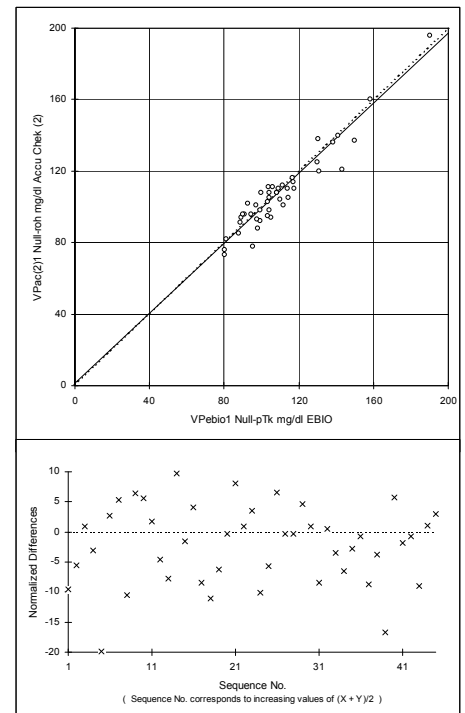
P/B Regression
 $Y = 1,100 * X - 1,719$
 $md(95) = 6,848$
 $N = 45, r = 0,969$

24 b: Nüchtern-Venenblut



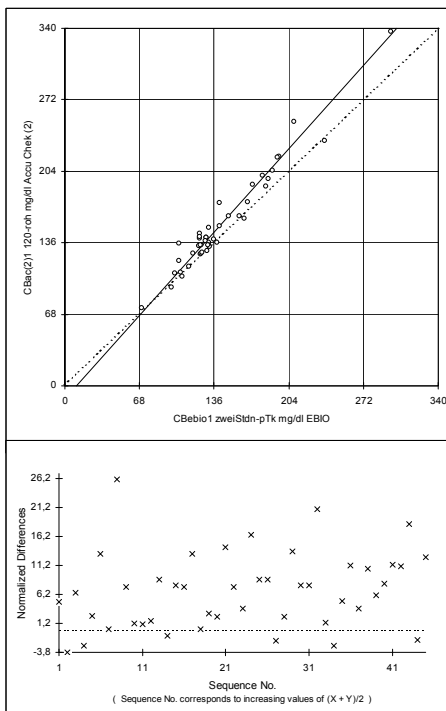
P/B Regression
 $Y = 1,084 * X - 0,842$
 $md(95) = 8,573$
 $N = 45, r = 0,952$

24 c: Nüchtern-Venenplasma



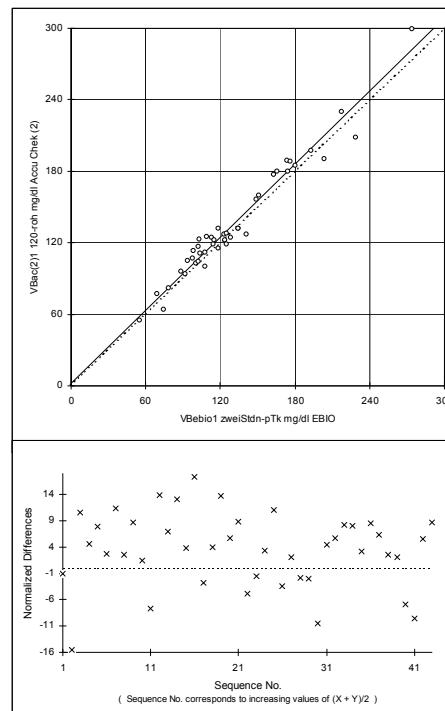
P/B Regression
 $Y = 0,979 * X + 1,605$
 $md(95) = 7,762$
 $N = 45, r = 0,950$

24 d: 2h-Postprandial-Kapillarblut



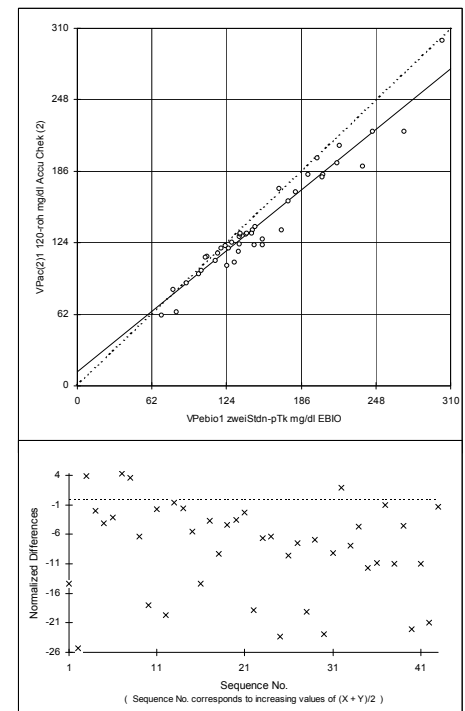
P/B Regression
 $Y = 1,165 * X - 11,936$
 $md(95) = 13,328$
 $N = 45, r = 0,976$

24 e: 2h-Postprandial-Venenblut



P/B Regression
 $Y = 1,023 * X + 1,963$
 $md(95) = 12,558$
 $N = 43, r = 0,980$

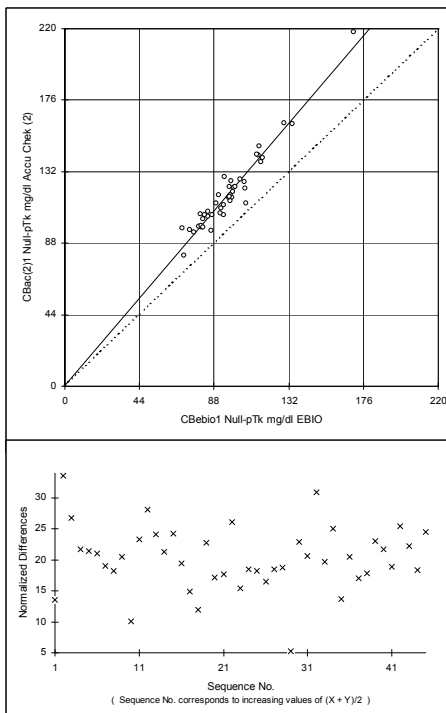
24 f: 2h-Postprandial-Venenplasma



P/B Regression
 $Y = 0,847 * X + 12,116$
 $md(95) = 16,298$
 $N = 43, r = 0,971$

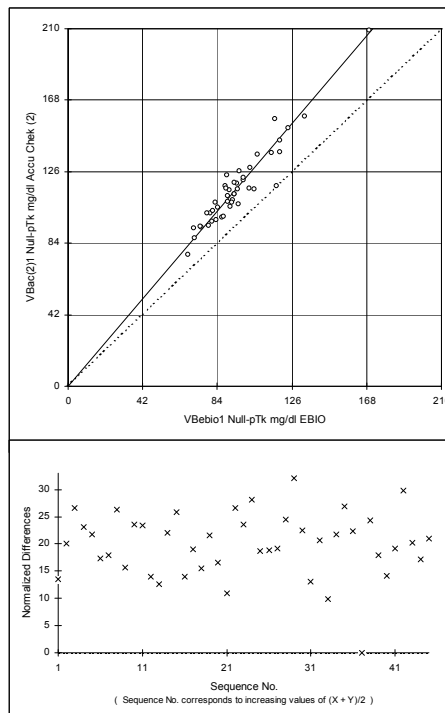
Abbildungen 25 a – f: Tages-korrigierte Vergleiche von Accu-Chek2 vs. Ebio

25 a: Nüchtern-Kapillarblut



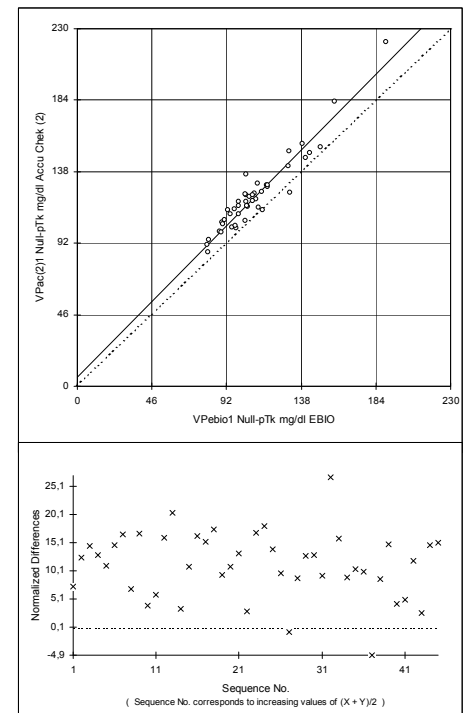
P/B Regression
 $Y = 1,222 * X + 0,475$
 $md(95) = 7,138$
 $N = 45, r = 0,967$

25 b: Nüchtern-Venenblut



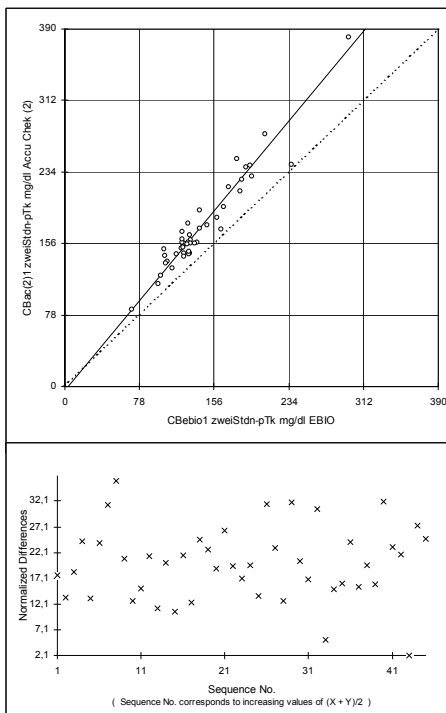
P/B Regression
 $Y = 1,226 * X - 0,075$
 $md(95) = 8,439$
 $N = 45, r = 0,950$

25 c: Nüchtern-Venenplasma



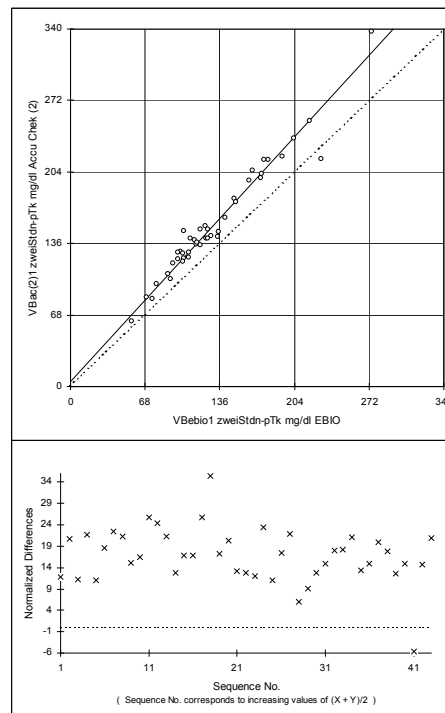
P/B Regression
 $Y = 1,060 * X + 5,835$
 $md(95) = 9,391$
 $N = 45, r = 0,953$

25 d: 2h-Postprandial-Kapillarblut



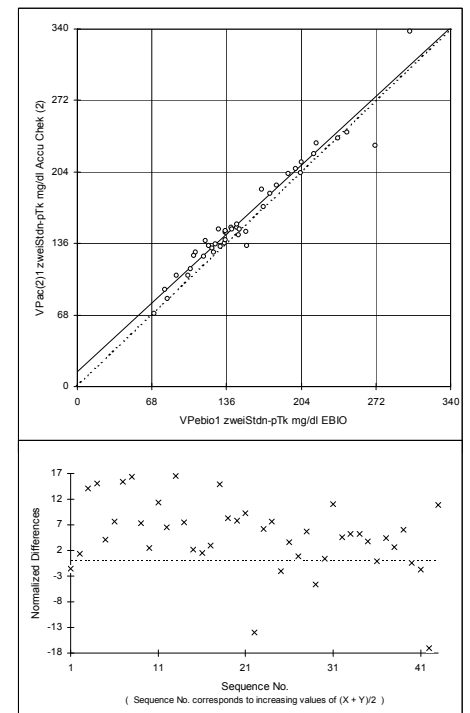
P/B Regression
 $Y = 1,258 * X - 4,496$
 $md(95) = 15,537$
 $N = 45, r = 0,962$

25 e: 2h-Postprandial-Venenblut



P/B Regression
 $Y = 1,141 * X + 4,819$
 $md(95) = 11,848$
 $N = 43, r = 0,977$

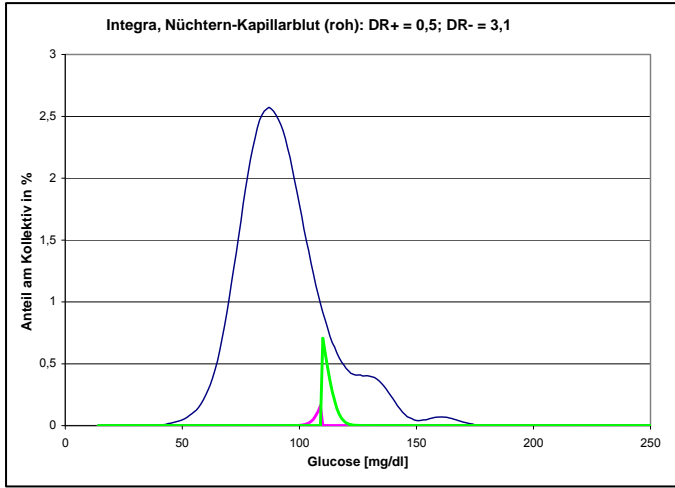
25 f: 2h-Postprandial-Venenplasma



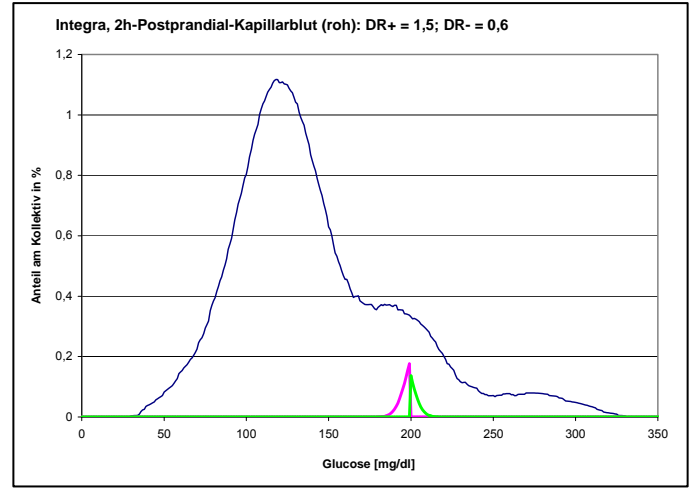
P/B Regression
 $Y = 0,961 * X + 14,270$
 $md(95) = 15,806$
 $N = 43, r = 0,973$

Abbildungen 28 a – f: Intersystemische unkorrigierte Vergleiche von Integra vs. Ebio. Graphische Darstellung der Ergebnisse des Diskordanzratentests.

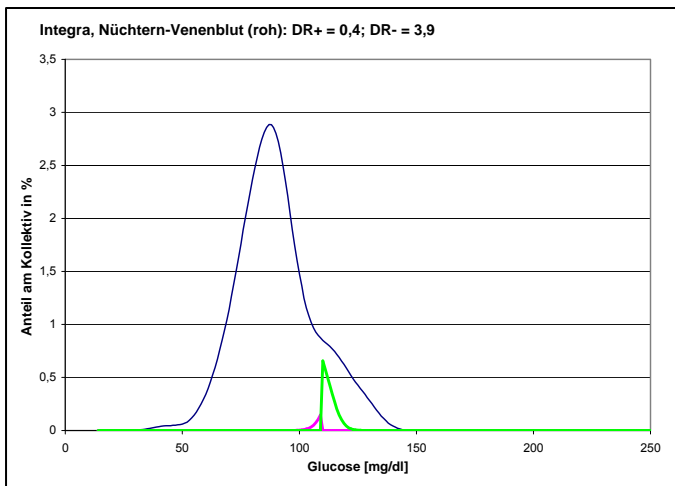
28 a:



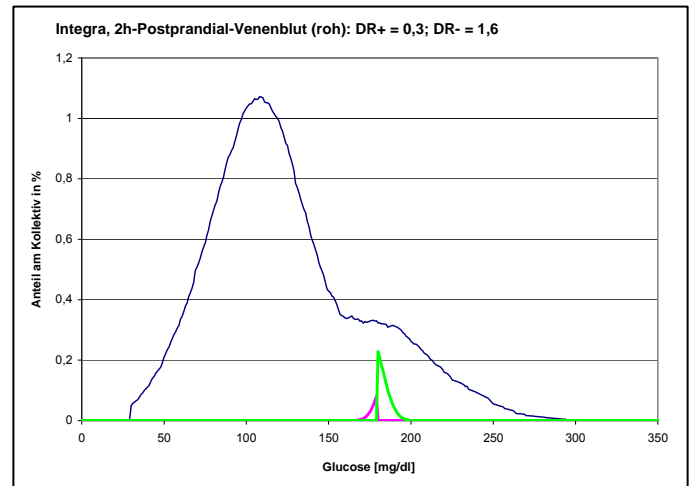
28 d:



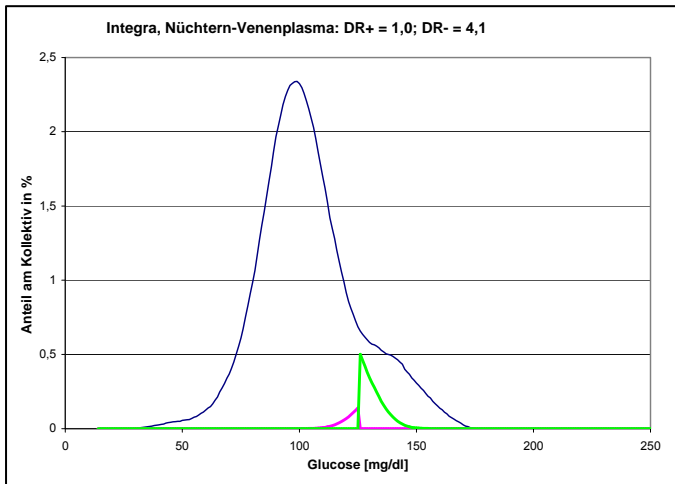
28 b:



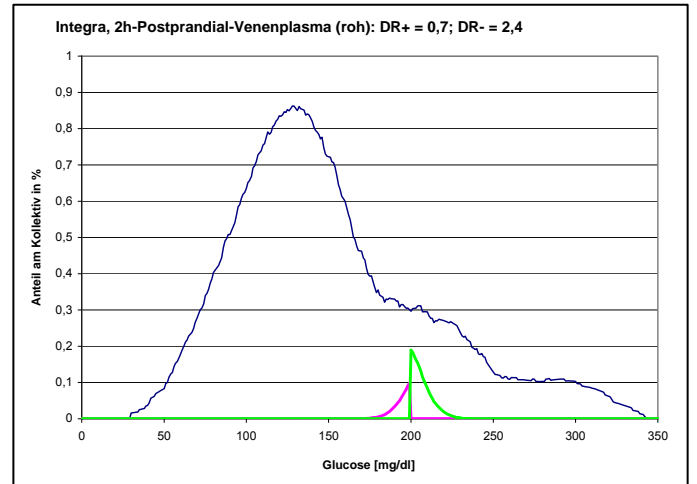
28 e:



28 c:

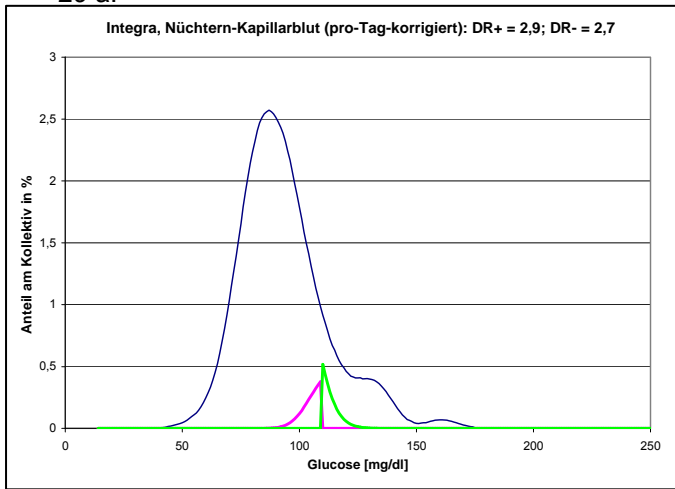


28 f:

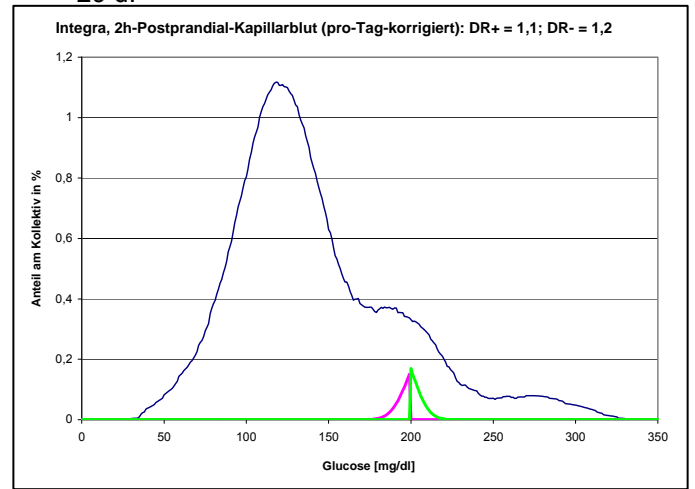


Abbildungen 29 a – f: Intersystemische Tages-korrigierte Vergleiche von Integra vs. Ebio. Graphische Darstellung der Ergebnisse des Diskordanzratentests

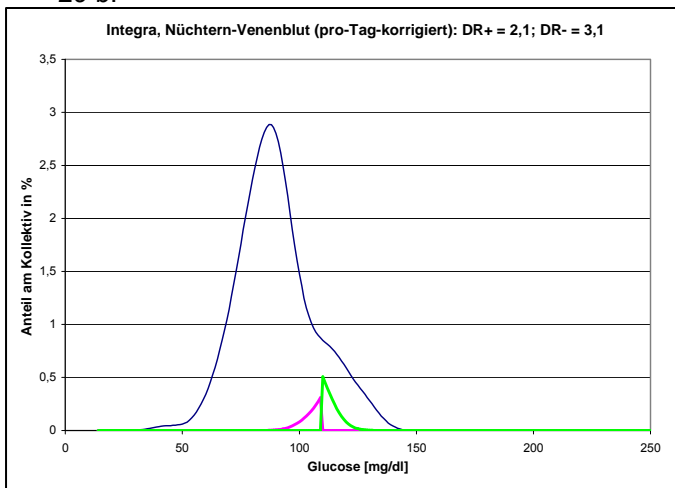
29 a:



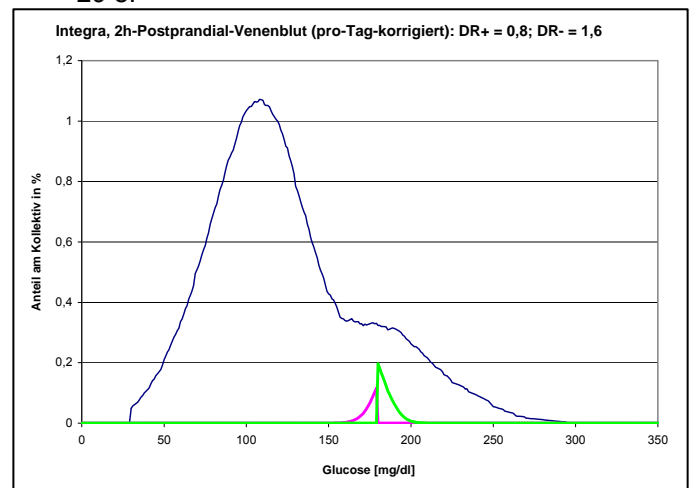
29 d:



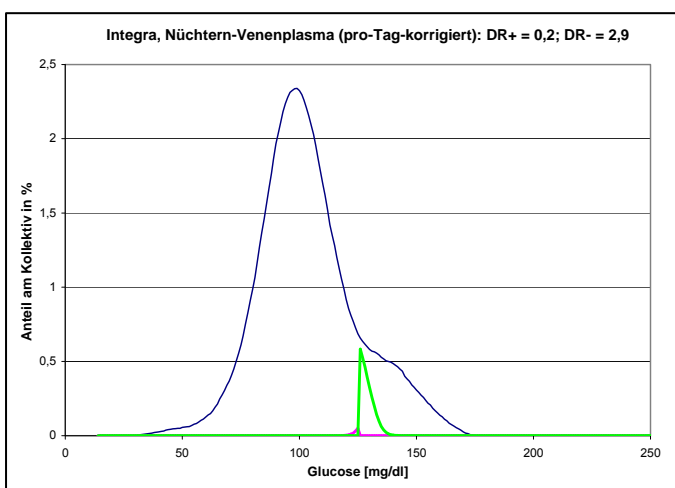
29 b:



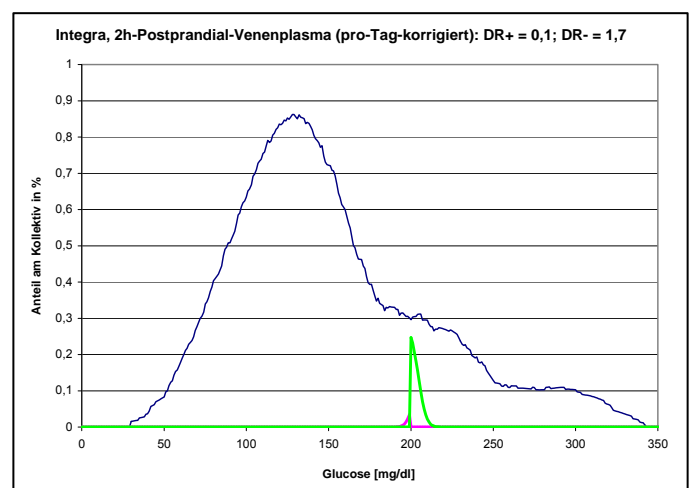
29 e:



29 c:

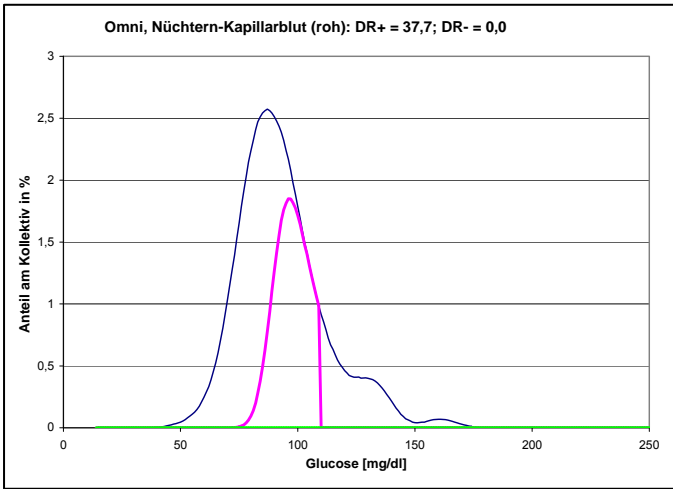


29 f:

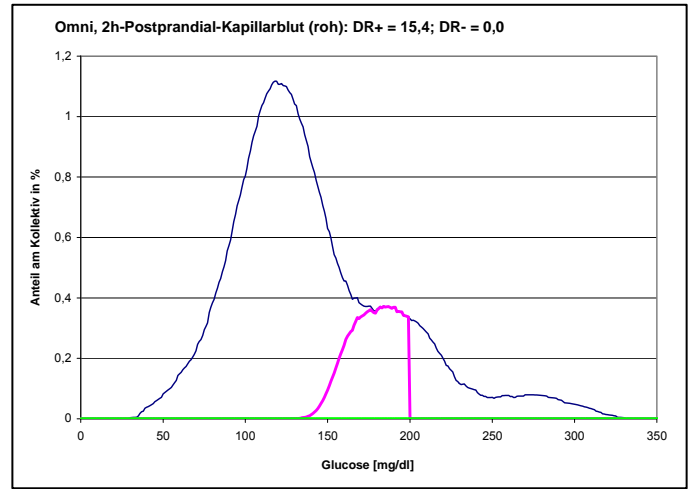


Abbildungen 30 a – f: Intersystemische unkorrigierte Vergleiche von Omni vs. Ebio. Graphische Darstellung der Ergebnisse des Diskordanzratentests.

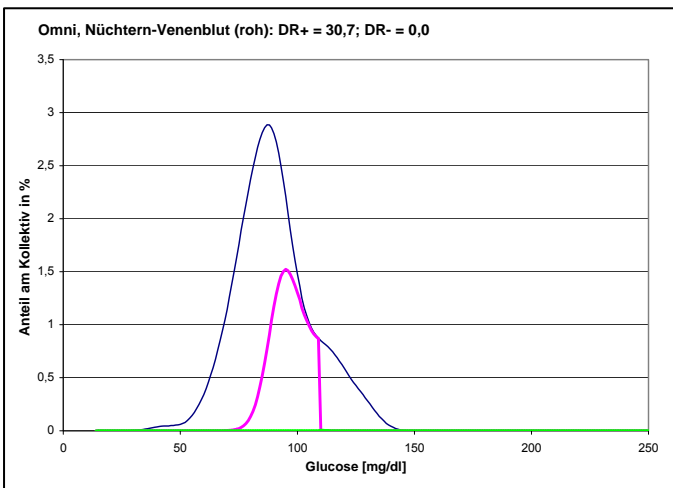
30 a:



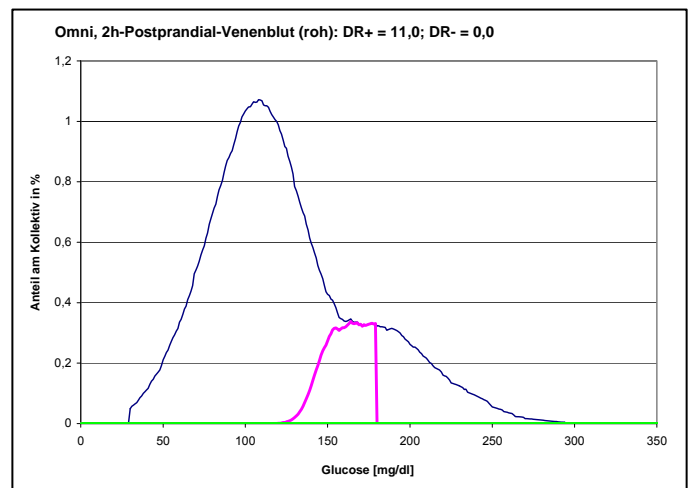
30 d:



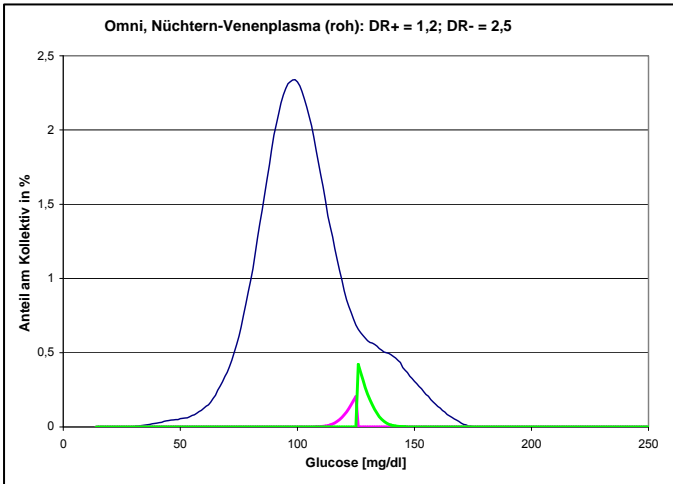
30 b:



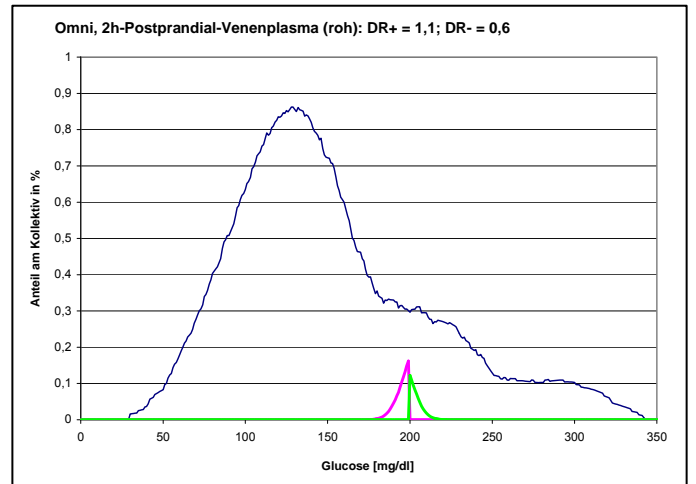
30 e:



30 c:

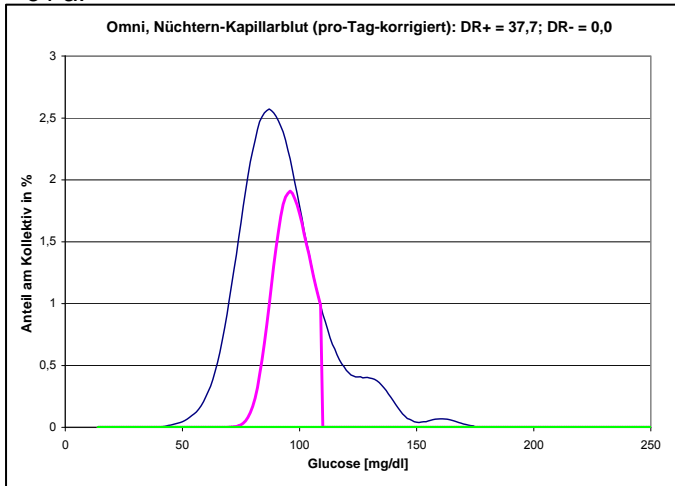


30 f:

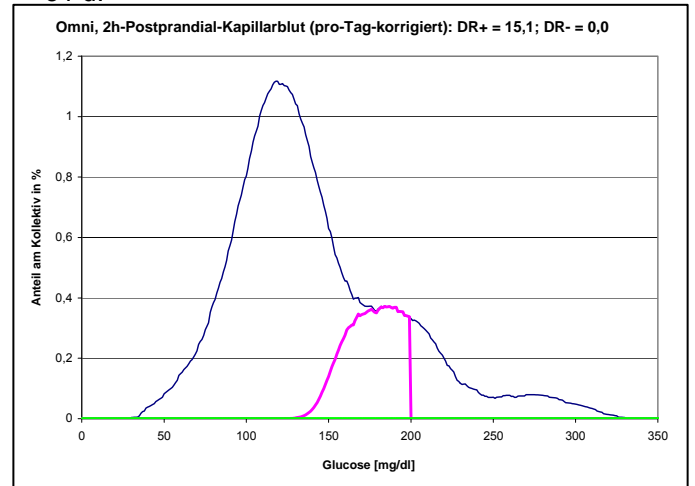


Abbildungen 31 a – f: Intersystemische Tages-korrigierte Vergleiche von Omni vs. Ebio. Graphische Darstellung der Ergebnisse des Diskordanzratentests

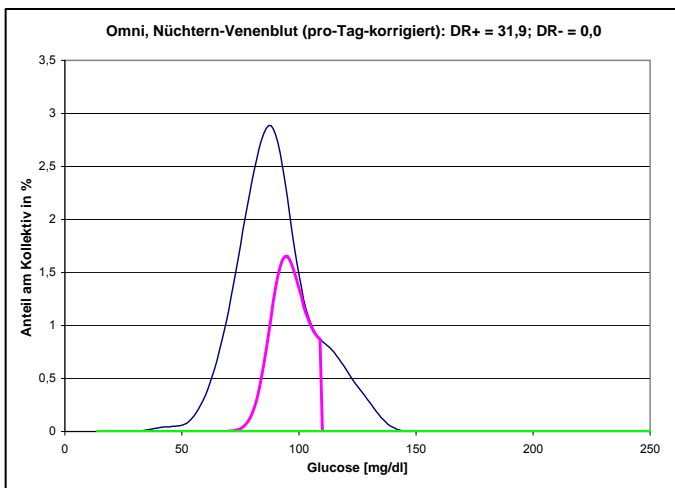
31 a:



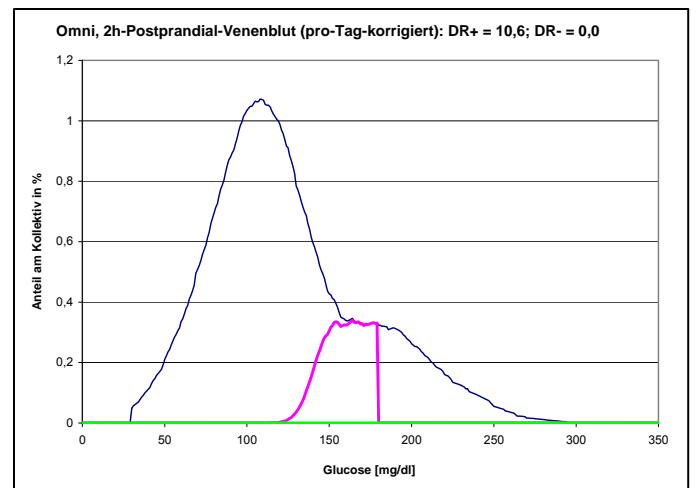
31 d:



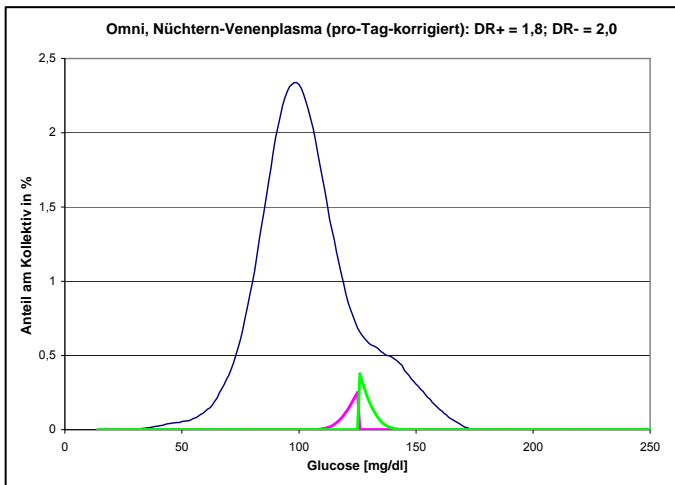
31 b:



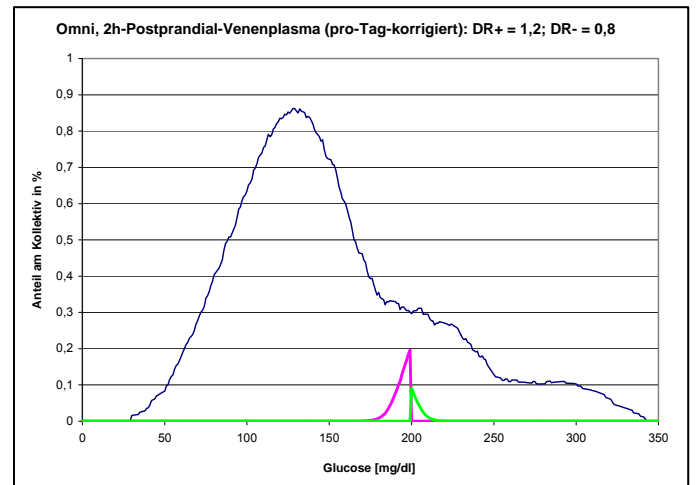
31 e:



31 c:

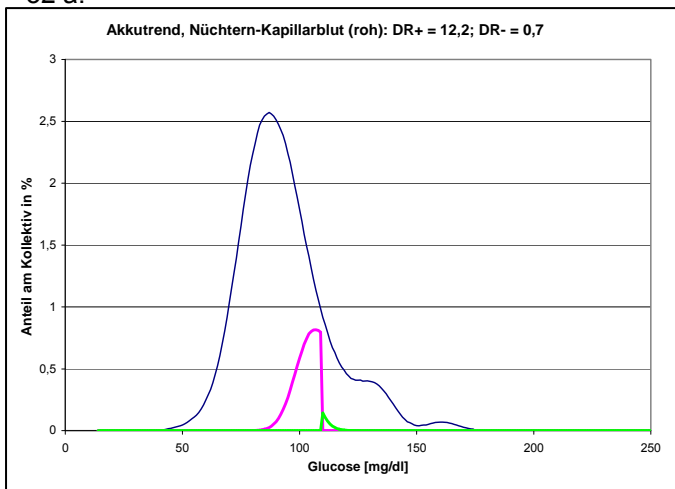


31 f:

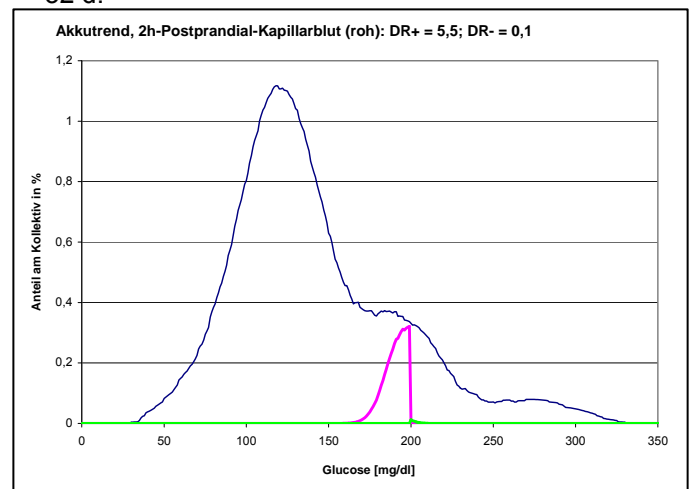


Abbildungen 32 a – f: Intersystemische unkorrigierte Vergleiche von Akkutrend vs. Ebio. Graphische Darstellung der Ergebnisse des Diskordanzratentests.

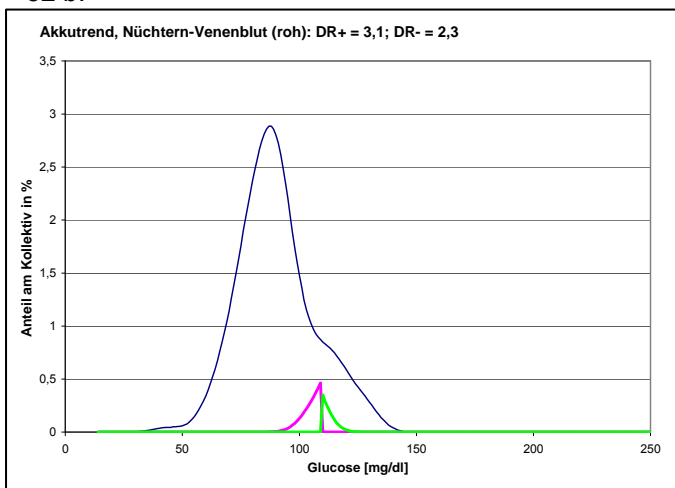
32 a:



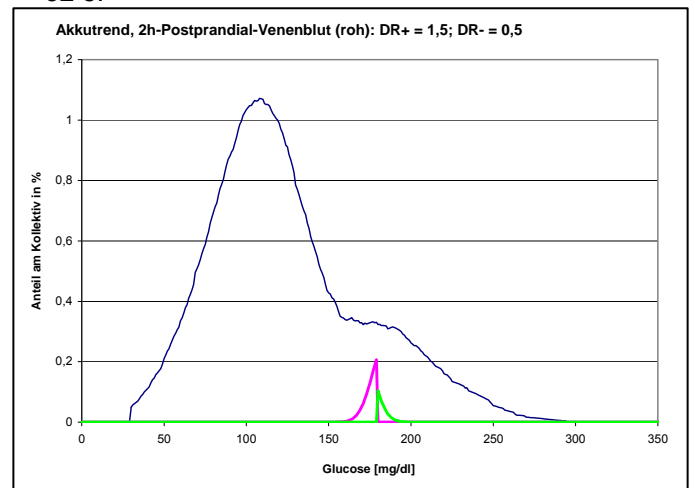
32 d:



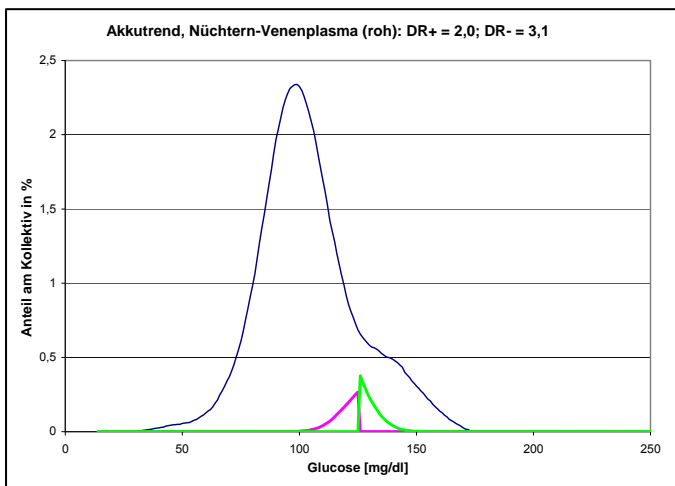
32 b:



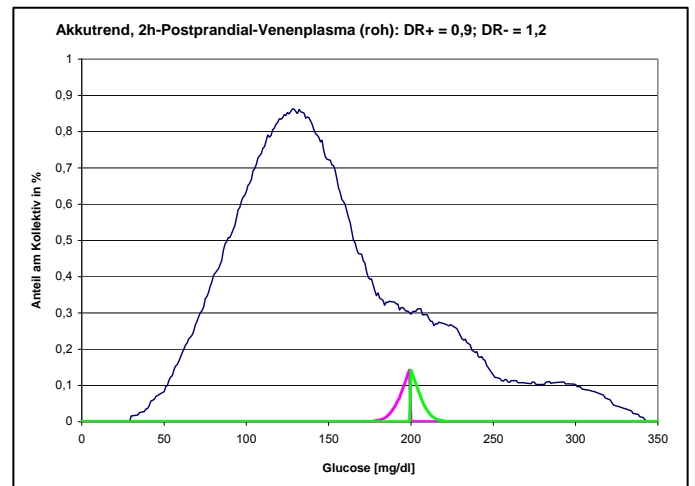
32 e:



32 c:

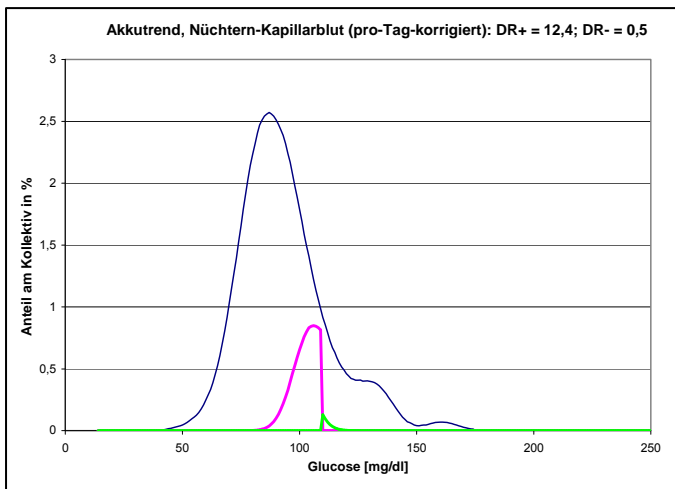


32 f:

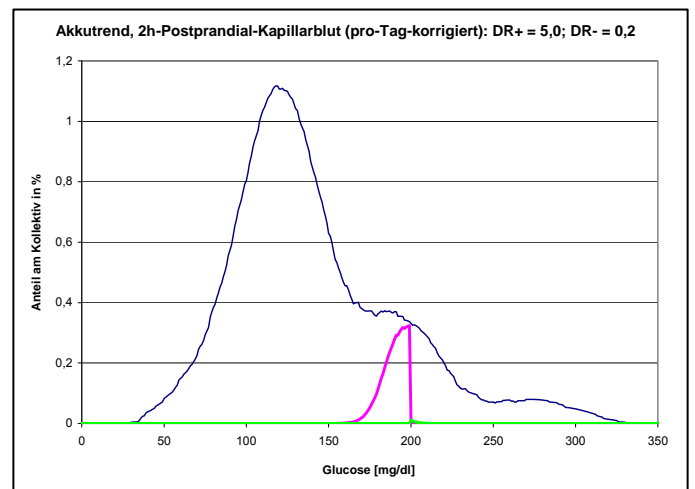


Abbildungen 33 a – f: Intersystemische Tages-korrigierte Vergleiche von Akkutrend vs. Ebio. Graphische Darstellung der Ergebnisse des Diskordanzratentests

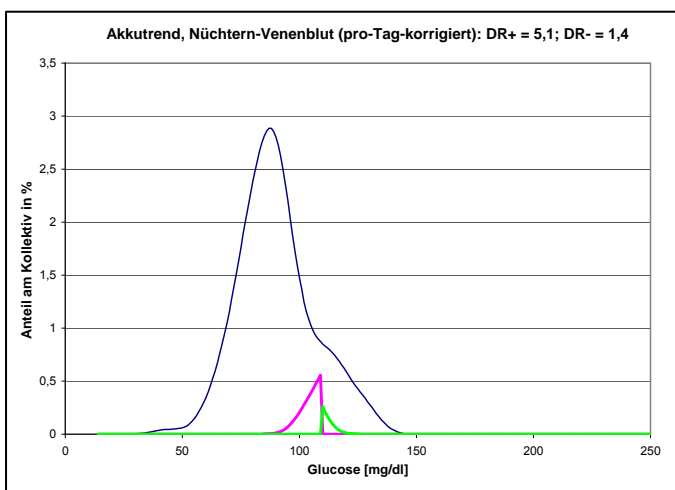
33 a:



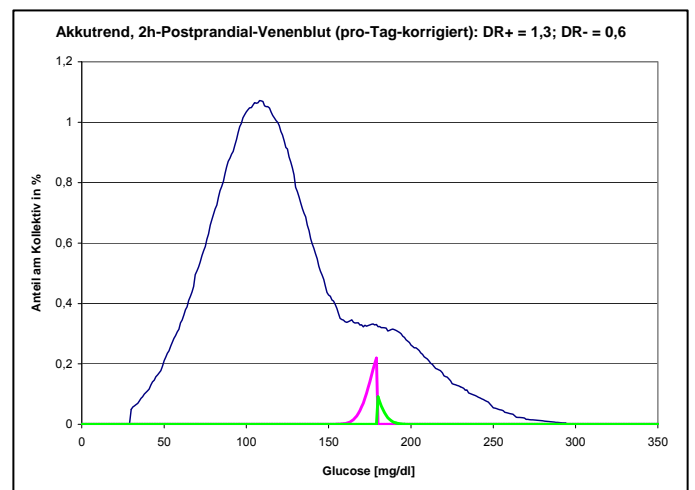
33 d:



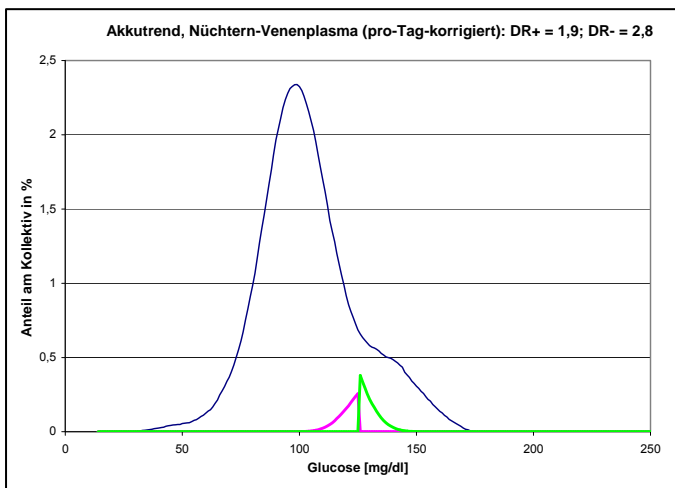
33 b:



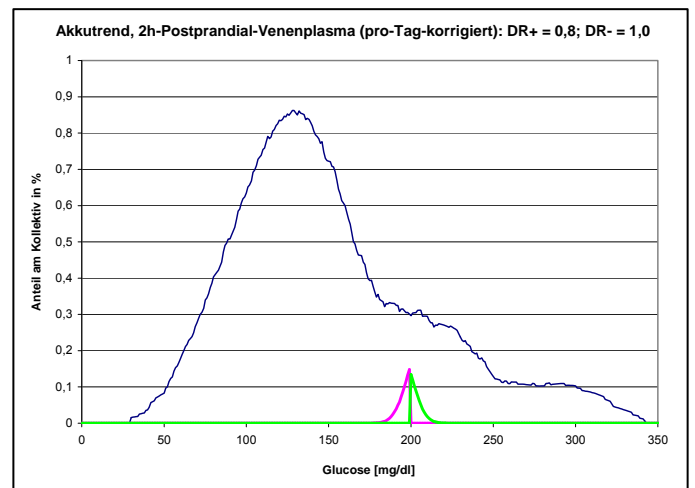
33 e:



33 c:

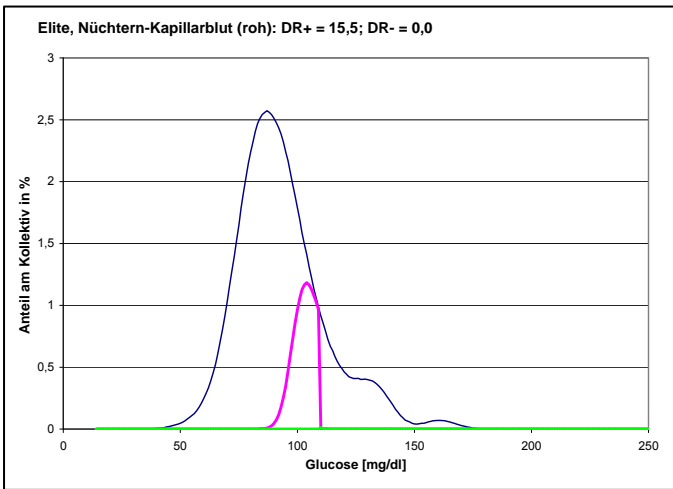


33 f:

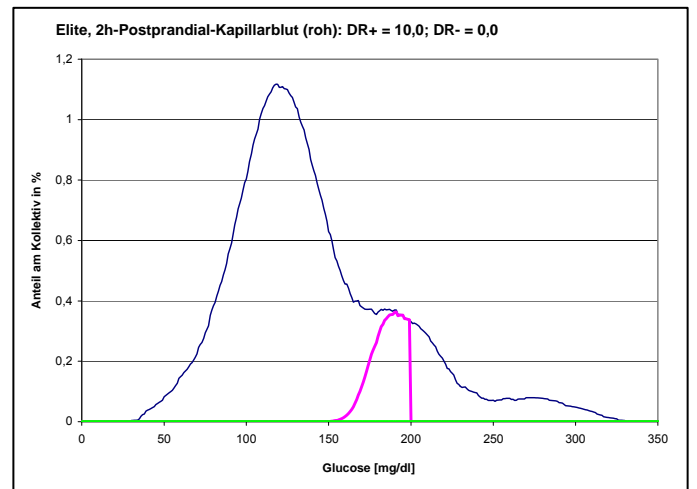


Abbildungen 34 a – f: Intersystemische unkorrigierte Vergleiche von Elite vs. Ebio. Graphische Darstellung der Ergebnisse des Diskordanzratentests.

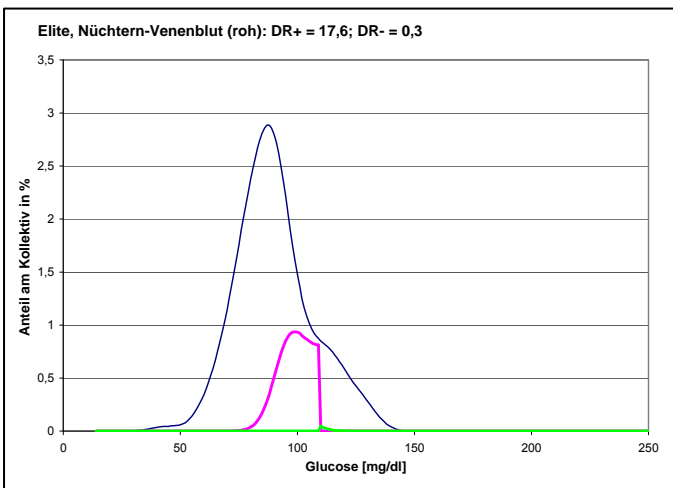
34 a:



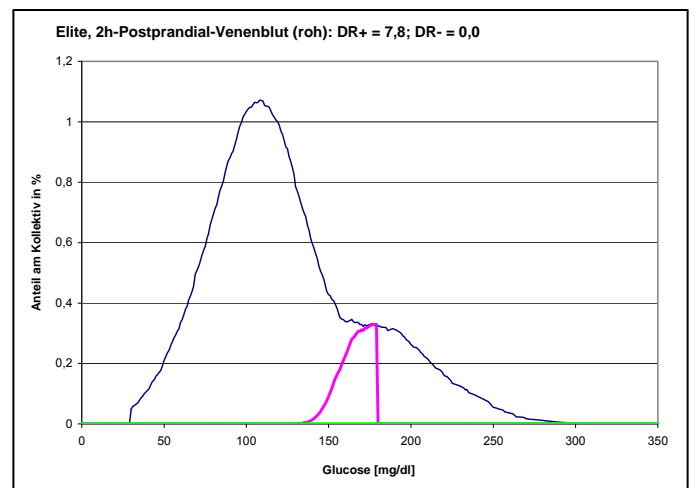
34 d:



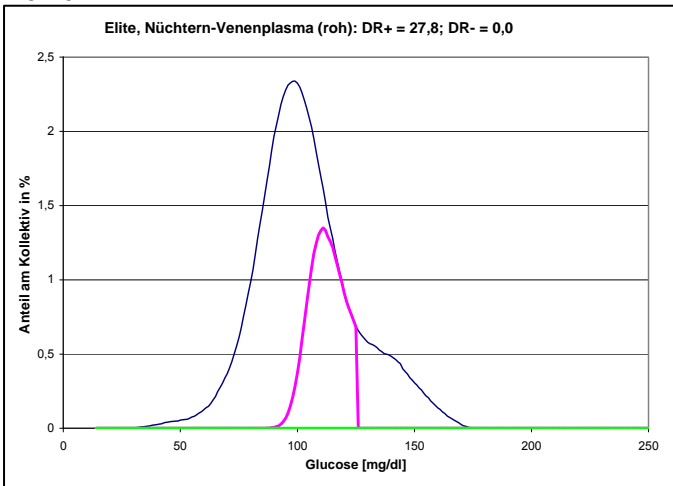
34 b:



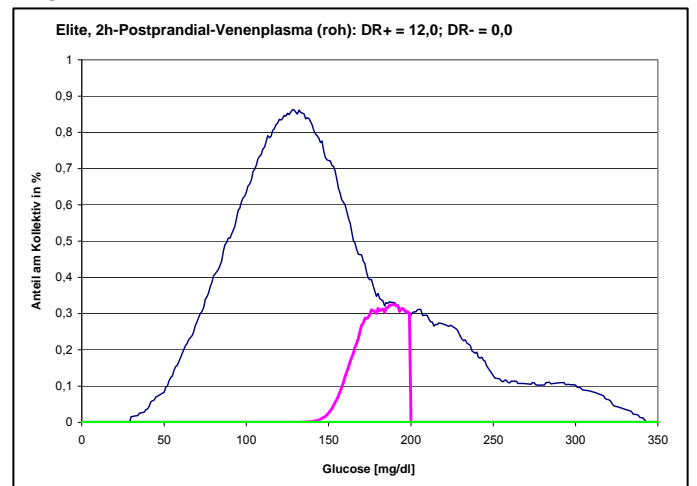
34 e:



34 c:

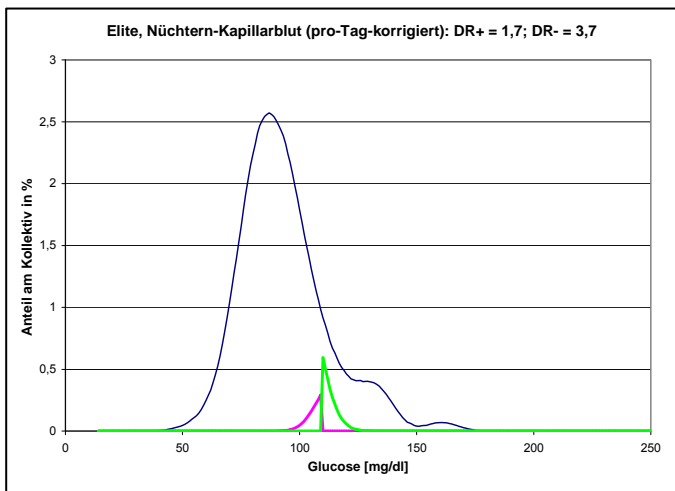


34 f:

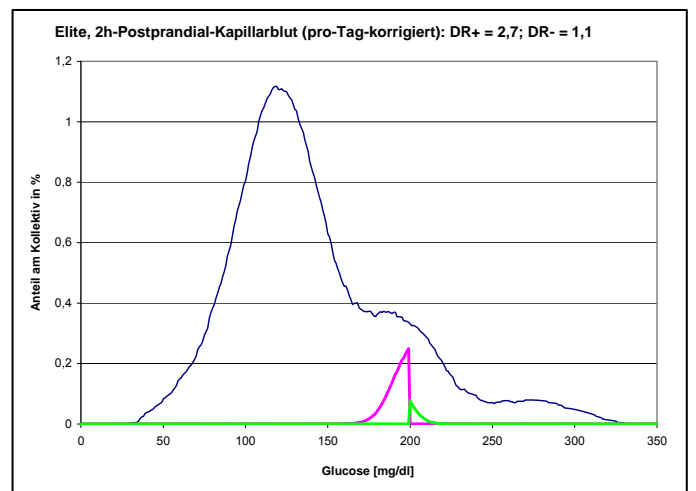


Abbildungen 35 a – f: Intersystemische Tages-korrigierte Vergleiche von Elite vs. Ebio. Graphische Darstellung der Ergebnisse des Diskordanzratentests

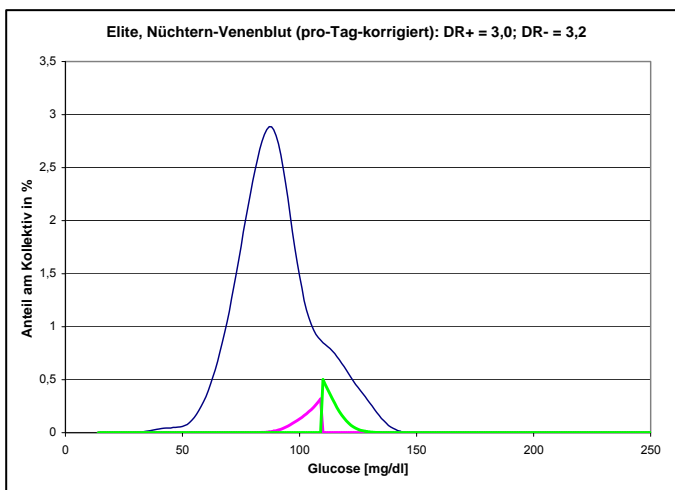
35 a:



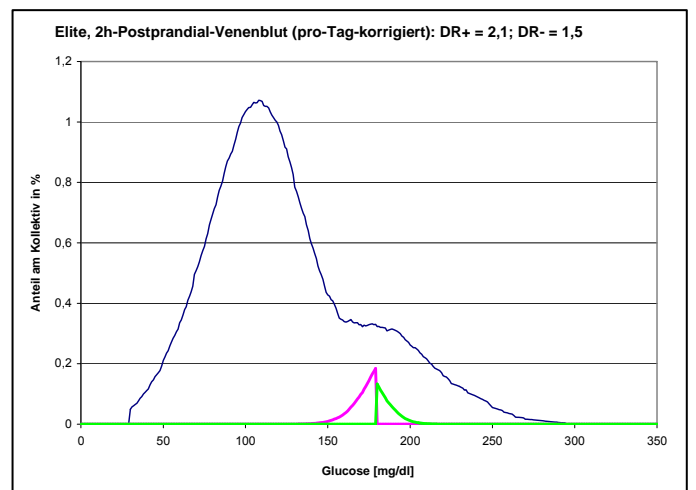
35 d:



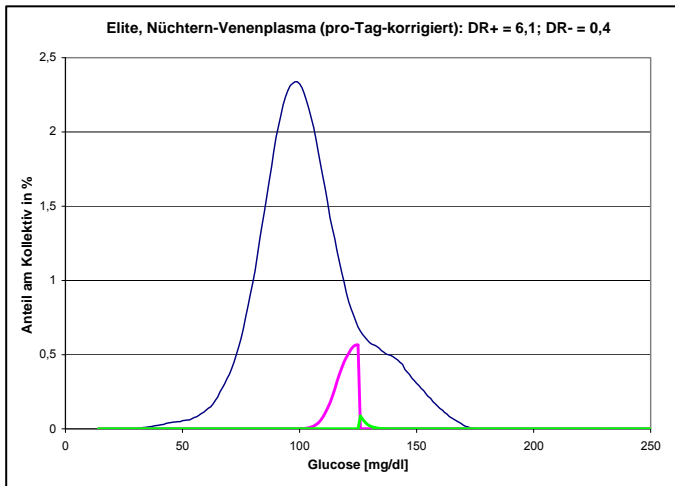
35 b:



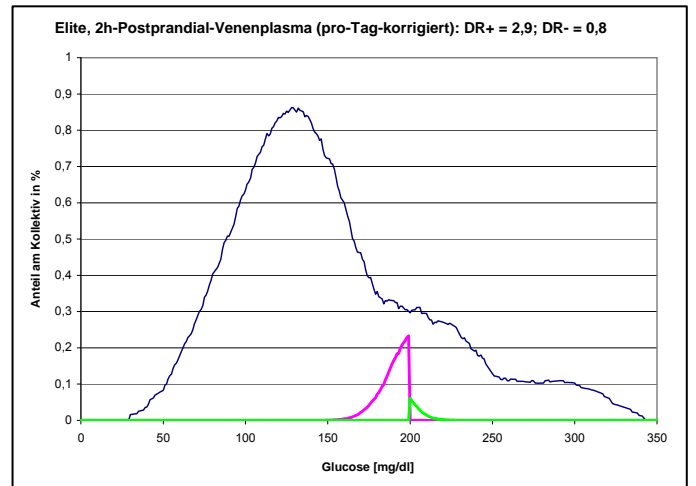
35 e:



35 c:

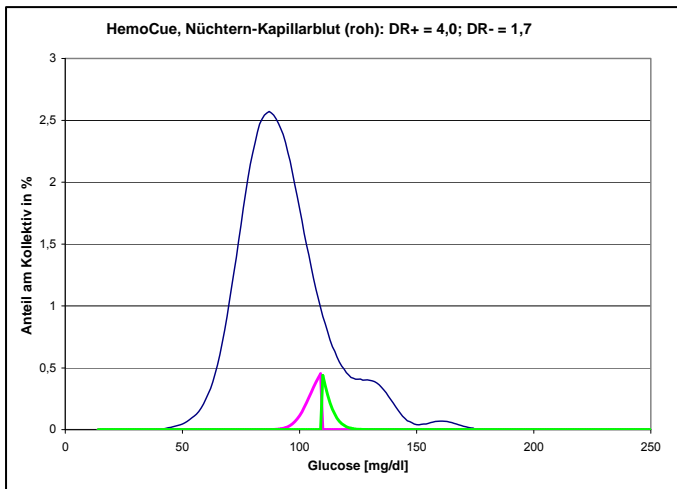


35 f:

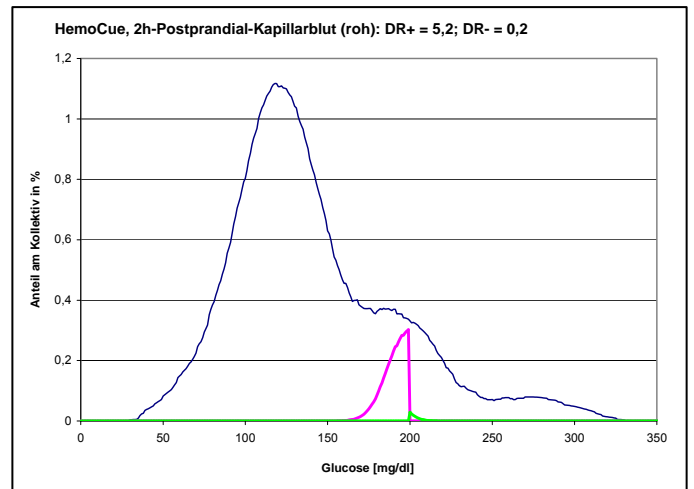


Abbildungen 36 a – f: Intersystemische unkorrigierte Vergleiche von HemoCue vs. Ebio. Graphische Darstellung der Ergebnisse des Diskordanzratentests.

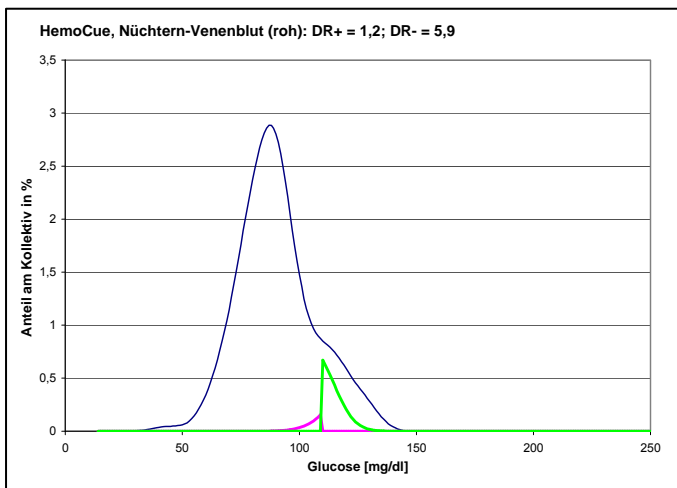
36 a:



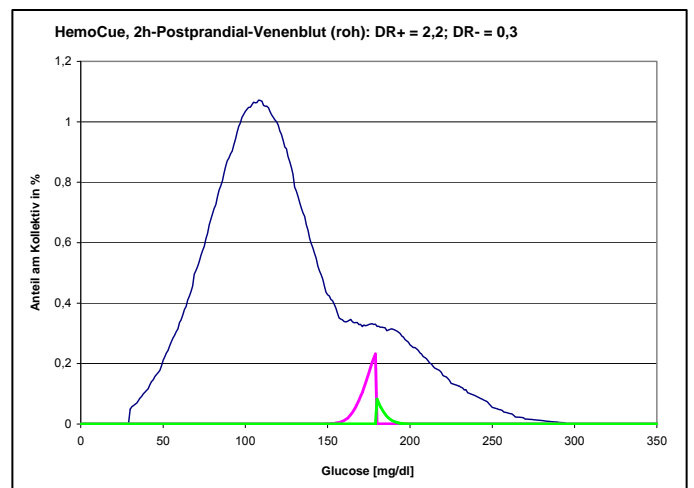
36 d:



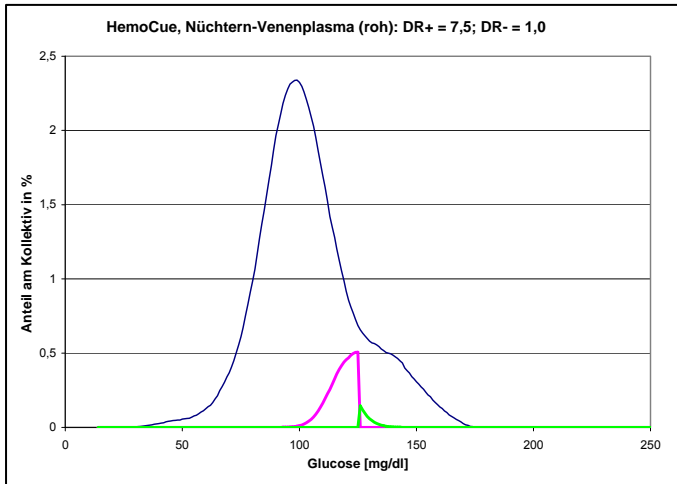
36 b:



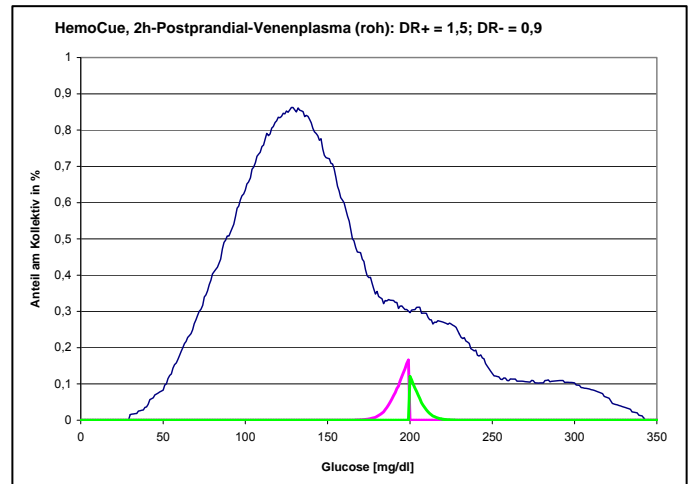
36 e:



36 c:

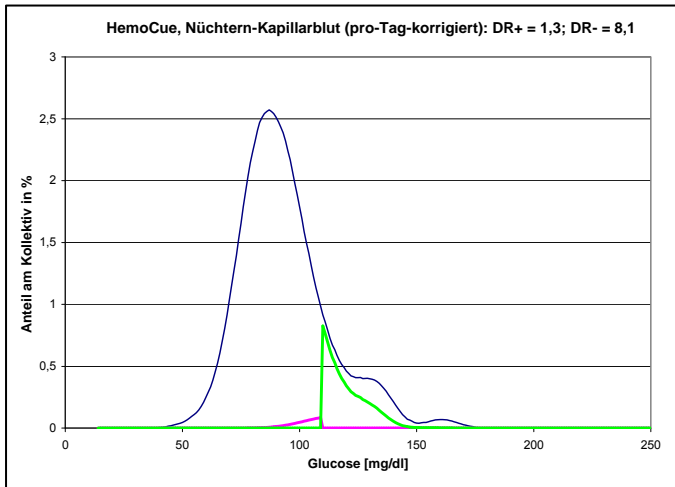


36 f:

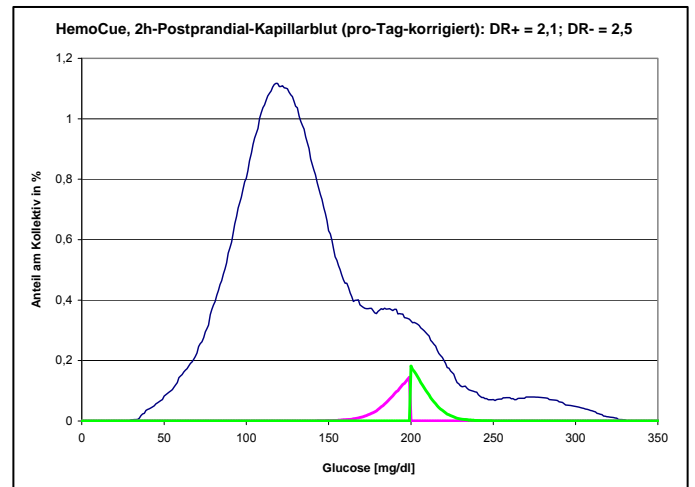


Abbildungen 37 a – f: Intersystemische Tages-korrigierte Vergleiche von HemoCue vs. Ebio. Graphische Darstellung der Ergebnisse des Diskordanzratentests

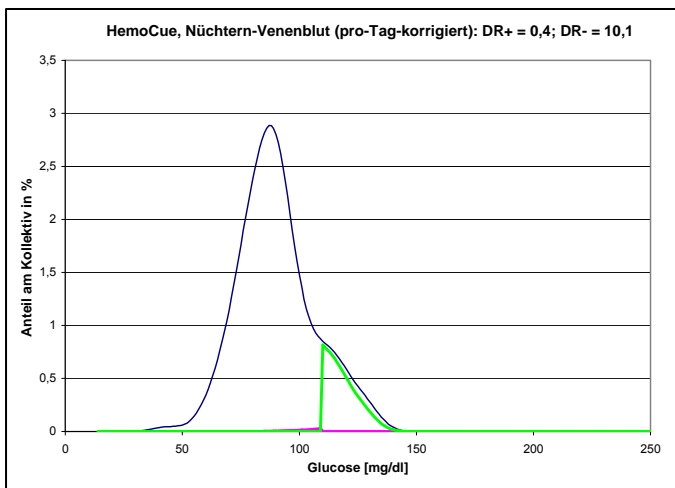
37 a:



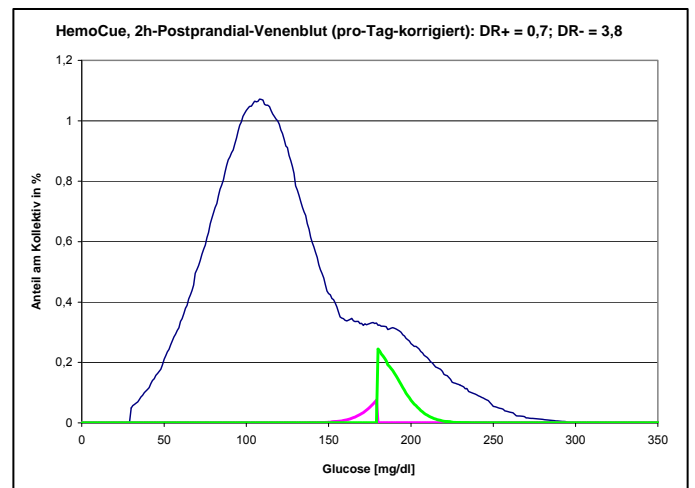
37 d:



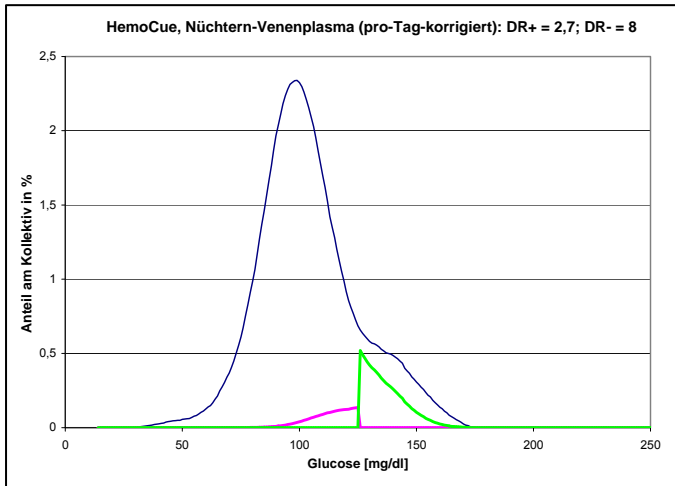
37 b:



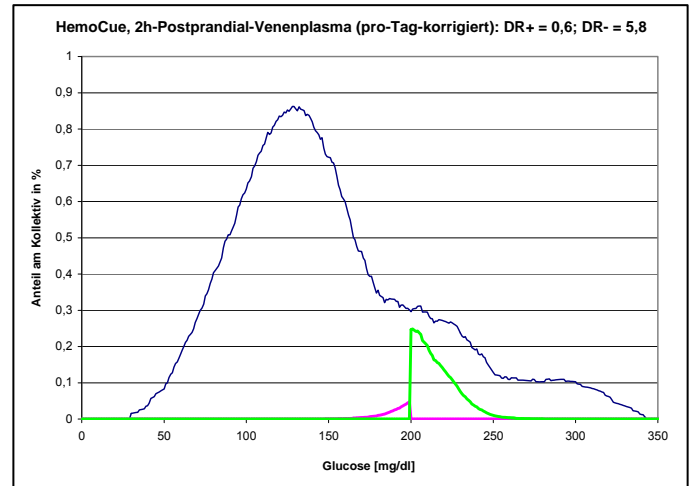
37 e:



37 c:

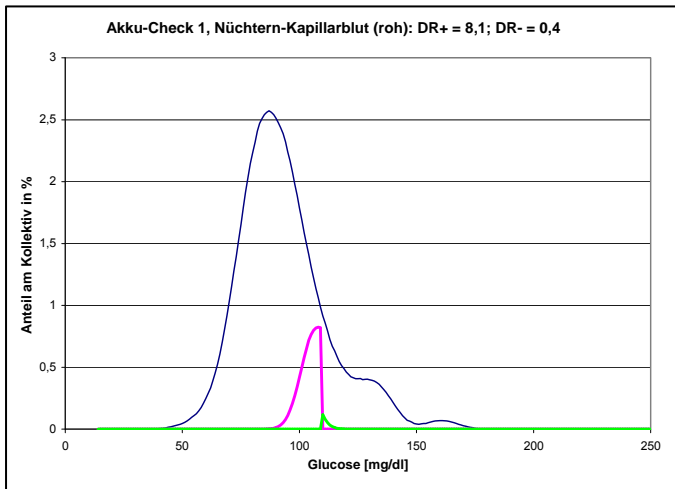


37 f:

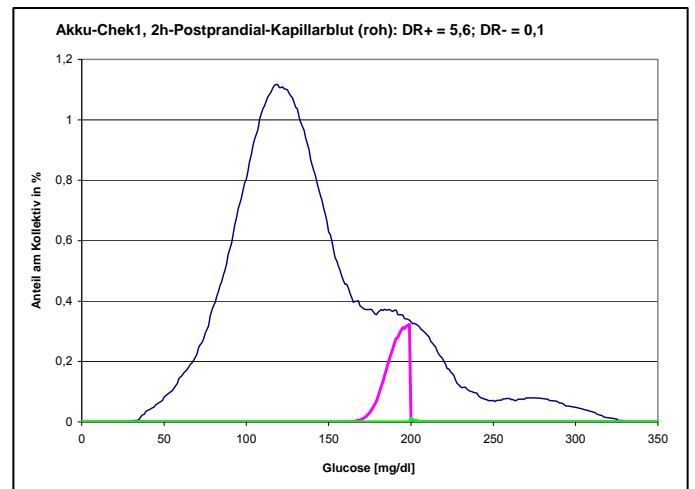


Abbildungen 38 a – f: Intersystemische unkorrigierte Vergleiche von Accu-Chek1 vs. Ebio. Graphische Darstellung der Ergebnisse des Diskordanzratentests.

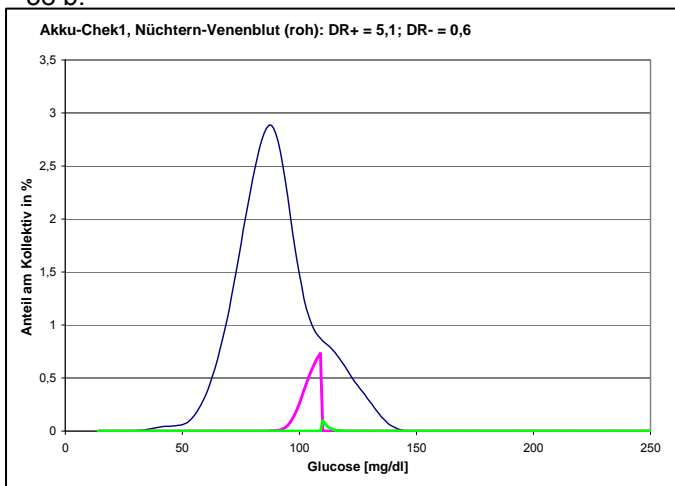
38 a:



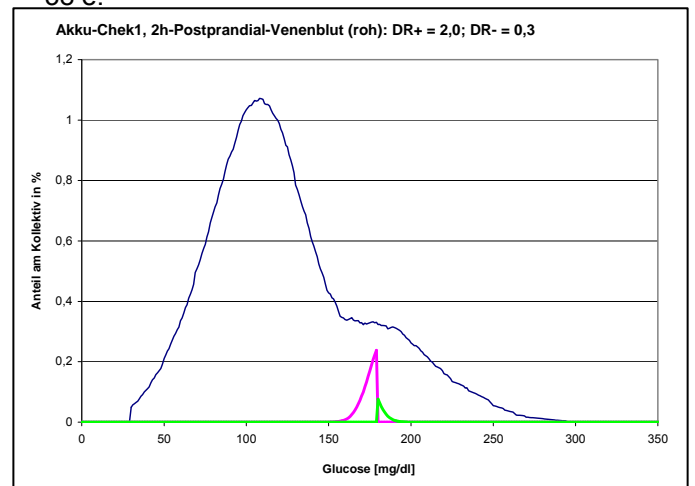
38 d:



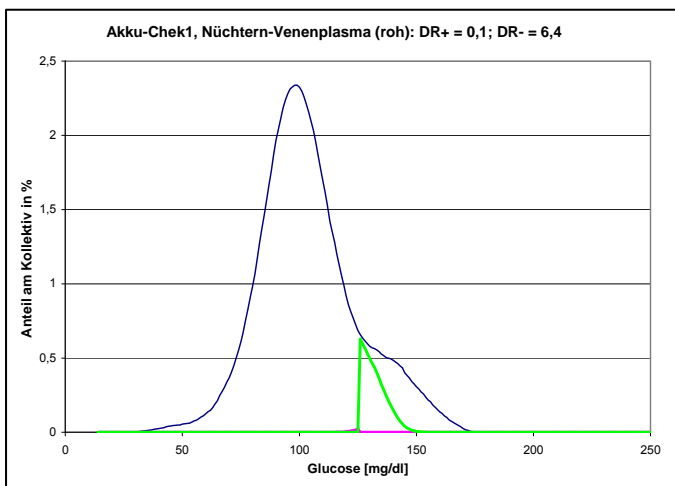
38 b:



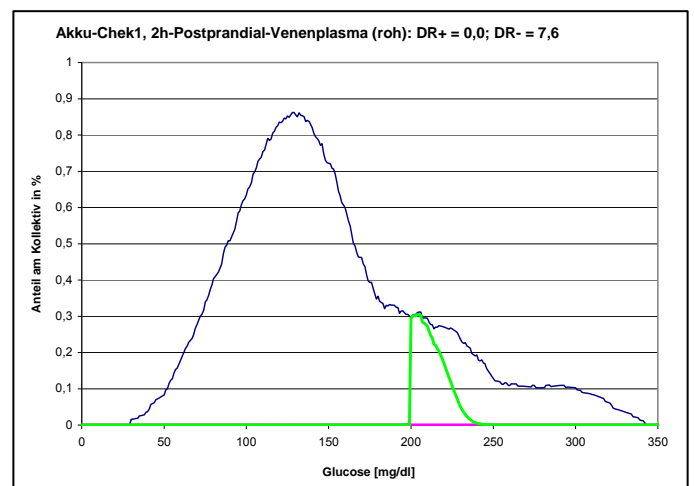
38 e:



38 c:

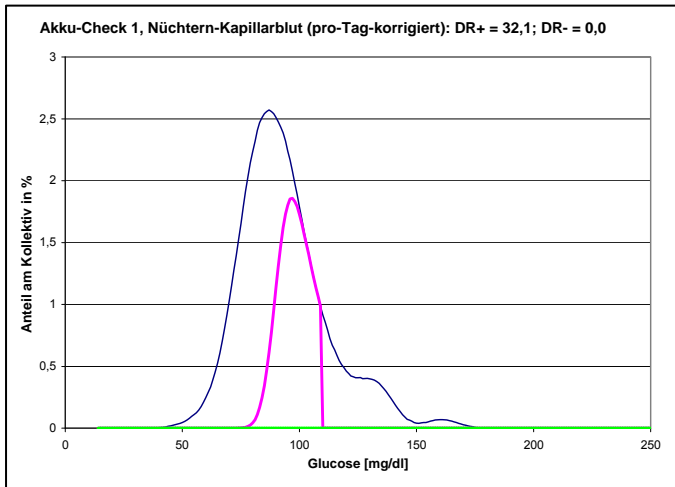


38 f:

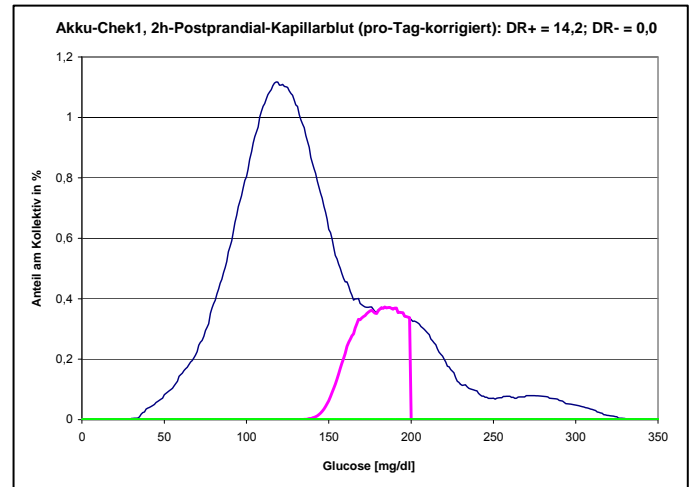


Abbildungen 39 a – f: Intersystemische Tages-korrigierte Vergleiche von Akku-Chek1 vs. Ebio. Graphische Darstellung der Ergebnisse des Diskordanzratentests

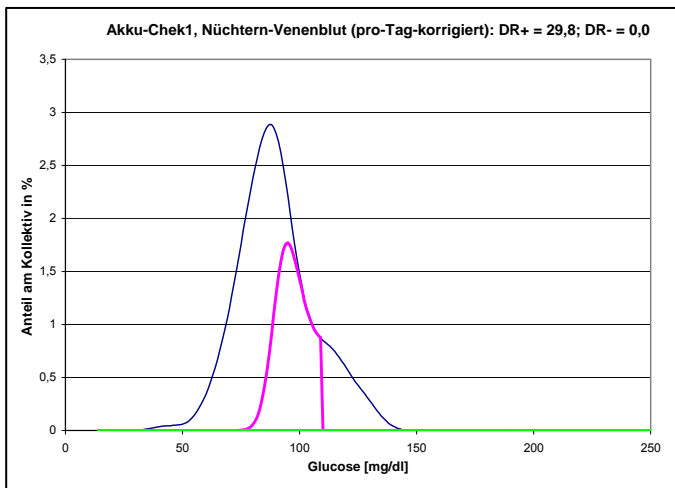
39 a:



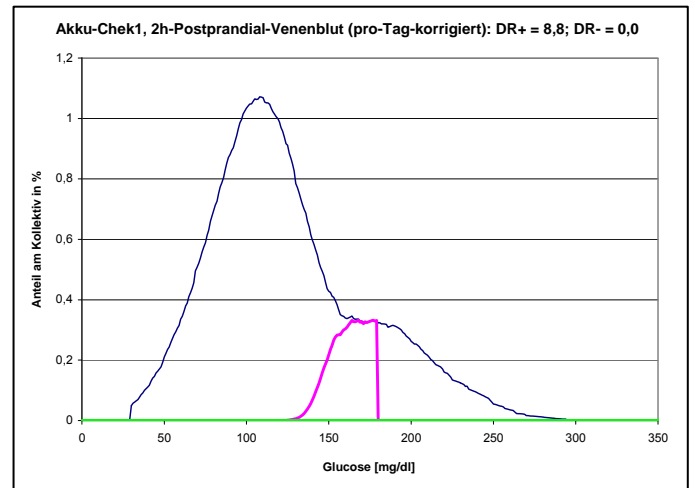
39 d:



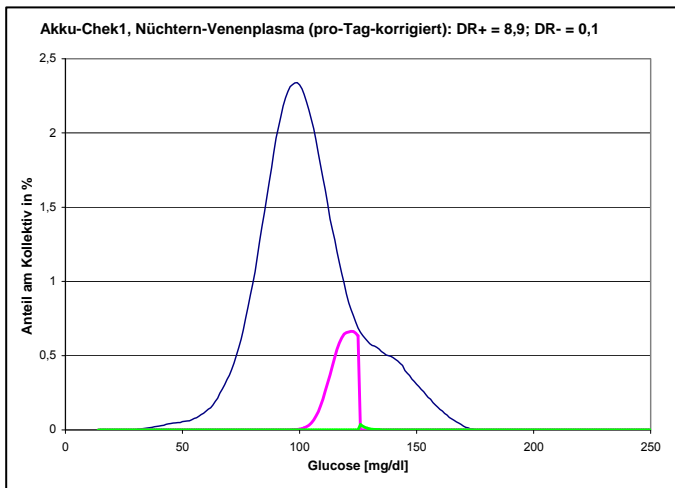
39 b:



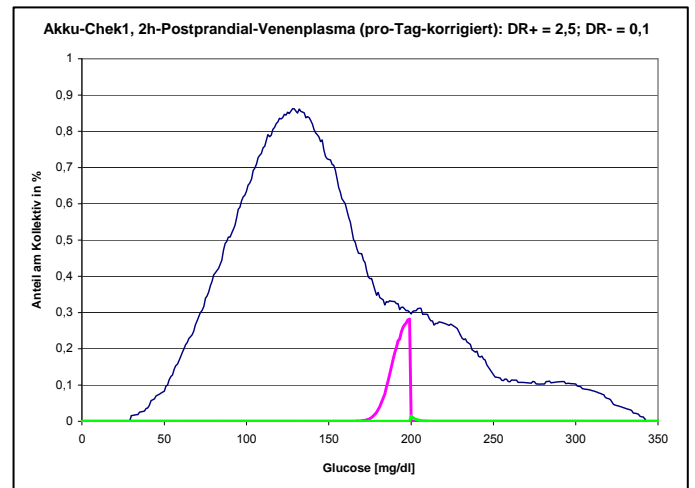
39 e:



39 c:

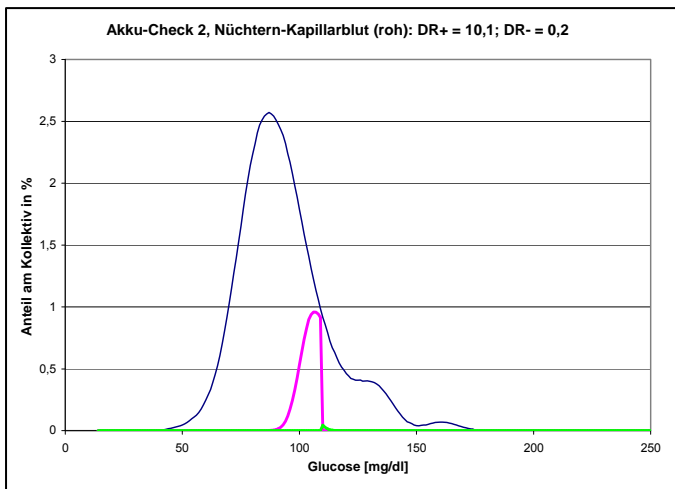


39 f:

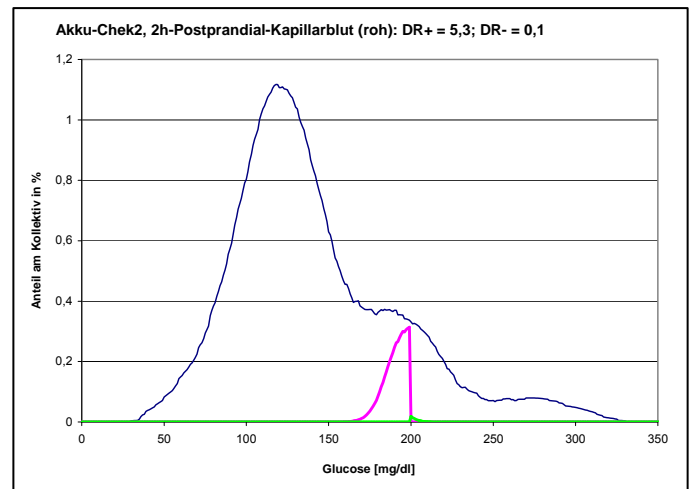


Abbildungen 40 a – f: Intersystemische unkorrigierte Vergleiche von Accu-Chek2 vs. Ebio. Graphische Darstellung der Ergebnisse des Diskordanzratentests.

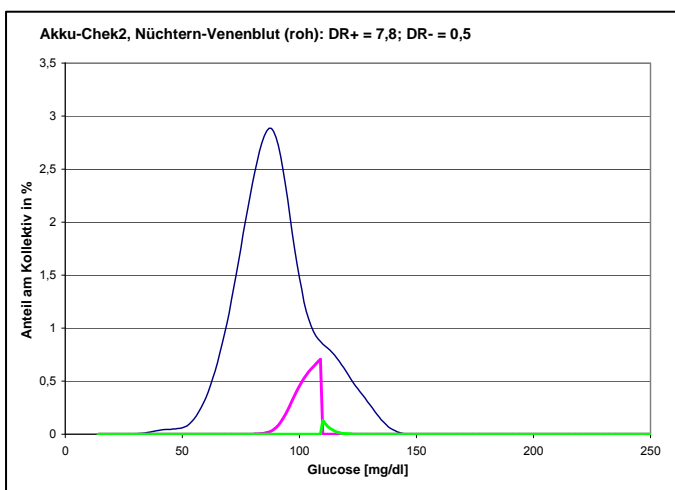
40 a:



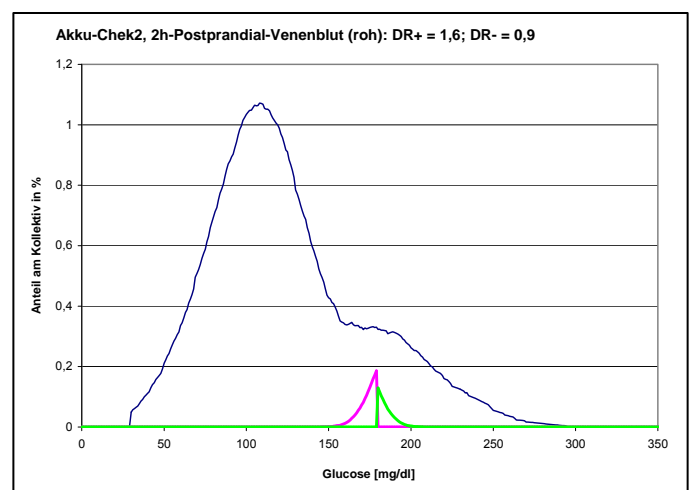
40 d:



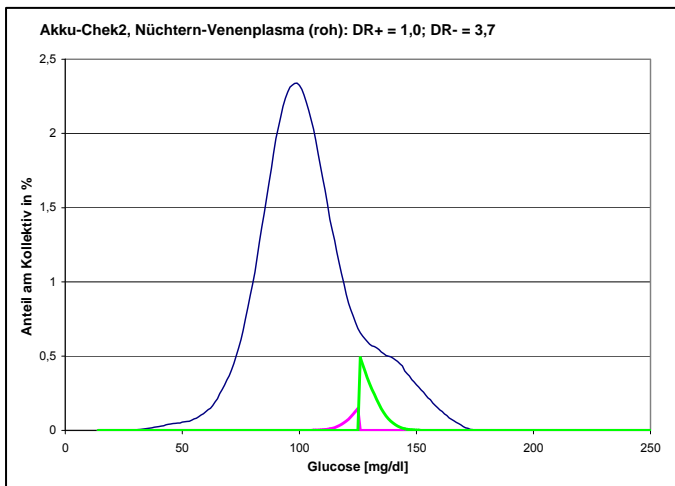
40 b:



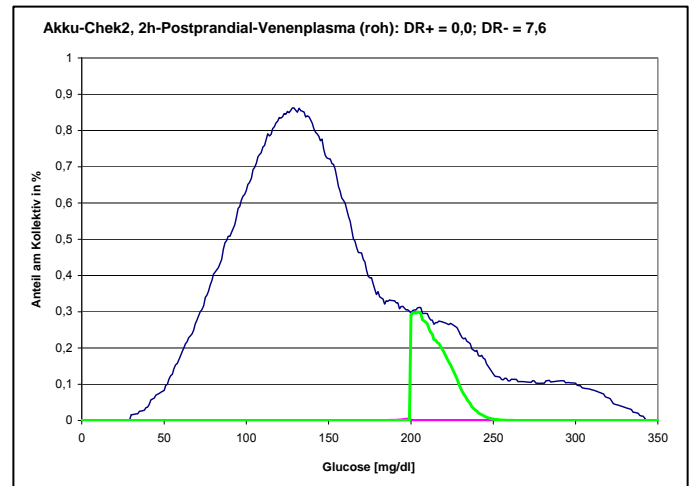
40 e:



40 c:

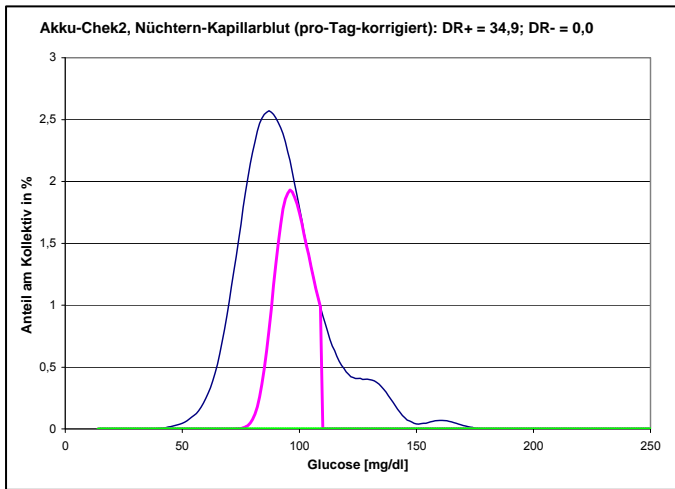


40 f:

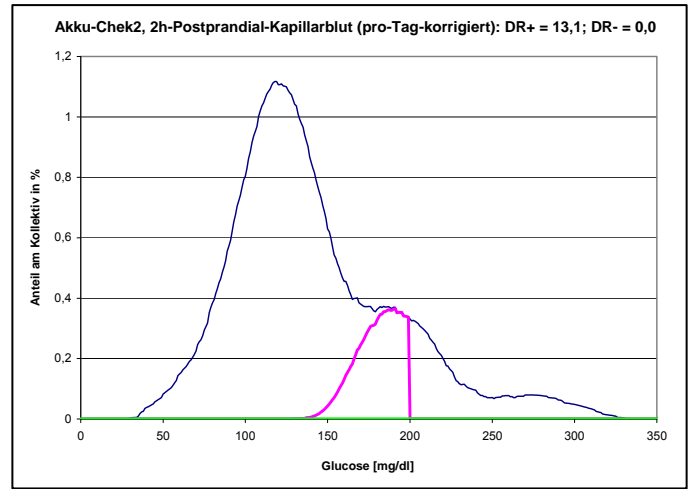


Abbildungen 41 a – f: Intersystemische Tages-korrigierte Vergleiche von Akku-Chek2 vs. Ebio. Graphische Darstellung der Ergebnisse des Diskordanzratentests

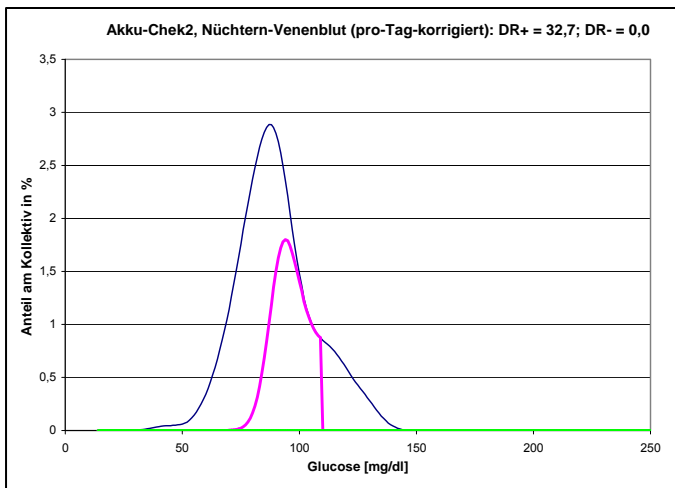
41 a:



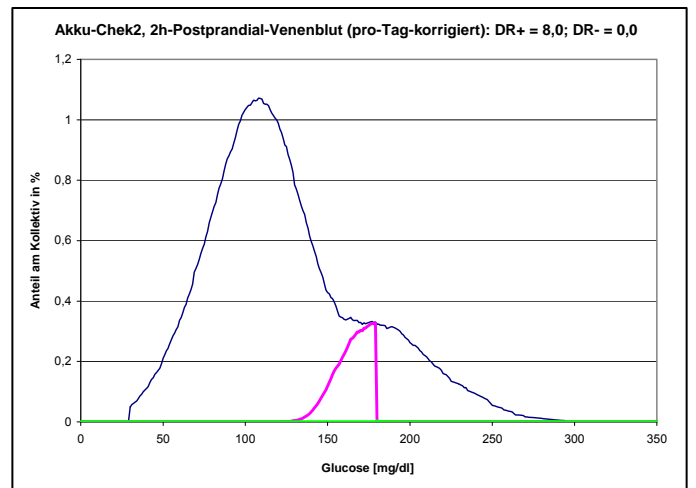
41 d:



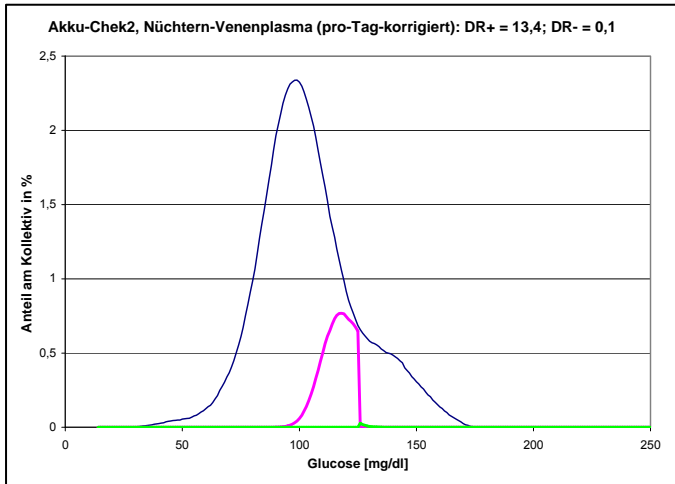
41 b:



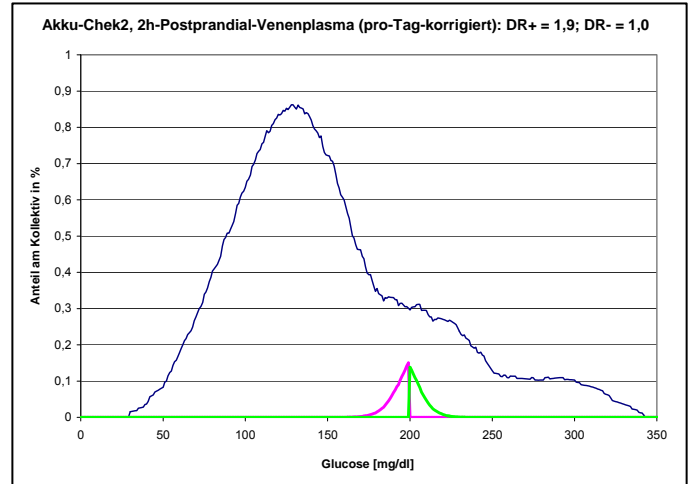
41 e:



41 c:



41 f:



8.3 TABELLENVERZEICHNIS

	Kapitel (erwähnt in Kapitel)
Tabelle 1: Maximal zulässige Fehlerrate für POCT-Glukometer, postuliert von Experten und Consensus-Gruppen	8.1 (1; 5)
Tabelle 2: "Hierarchy of Objectivity" (nach Fraser und Petersen)	8.1 (1; 5)
Tabelle 3: Gerätespezifikationen, Chargen der Teststreifen	3.1.1.2
Tabelle 4: Anwendungszeiträume und Einsatzfrequenz der Glukometer während des Studienkollektiv-Zeitraums	3.1.1.2
Tabelle 5: Materialien	3.1.2
Tabelle 6: Reagenzien	3.1.2
Tabelle 7: Glukosestandards und Kontroll-Lösungen	3.2.1
Tabelle 8: Diagnostische Kriterien hinsichtlich des Glukosestoffwechsels entsprechend der Empfehlungen der WHO 1999	8.1 (3.2.2.4)
Tabelle 9: Patientencharakteristika der Populationsprobanden	3.2.2.4
Tabelle 10: Anzahl an Messungen der Populationsprobanden pro Probenart	3.2.2.4
Tabelle 11: Kenngrößen diagnostischer Tests	8.1 (3.3.1; 5)
Tabelle 12: Messergebnisse in den Standard- und Kontroll-Lösungen	4.1.1
Tabelle 13: Tages-Korrekturfaktoren, ermittelt anhand des Glu-NIST	4.1.2
Tabelle 14: Anzahl an falsch positiven und falsch negativen Messungen bei den „intrasystemischen“ Vergleichen bezüglich der Entscheidungsgrenze „diabetisch / nicht-diabetisch,,	4.2.1
Tabelle 15: Anzahl an falsch positiven und falsch negativen Messungen bei den „intersystemischen“ Vergleichen bezüglich der Entscheidungsgrenze „diabetisch / nicht-diabetisch“	4.2.1
Tabelle 16: Intersystemische Vergleiche zum Nüchternzeitpunkt. Ergebnisse der linearen Regressionsanalysen nach Passing-Bablok	4.2.2
Tabelle 17: Intersystemische Vergleiche zum 2h-postprandialzeitpunkt. Ergebnisse der linearen Regressionsanalysen nach Passing-Bablok	4.2.2
Tabelle 18: Gesamt-Diskordanzraten (%) der intrasystemischen Vergleiche zum Nüchternzeitpunkt	4.2.3.1
Tabelle 19: Gesamt-Diskordanzraten (%) der intrasystemischen Vergleiche zum 2h-postprandialzeitpunkt	4.2.3.1

	Kapitel (erwähnt in Kapitel)
Tabelle 20: Gesamt-Diskordanzraten (%) der intersystemischen Vergleiche zum Nüchternzeitpunkt	4.2.3.2
Tabelle 21: Gesamt-Diskordanzraten (%) der intersystemischen Vergleiche zum 2h-postprandialzeitpunkt	4.2.3.2

8.4 ABBILDUNGSVERZEICHNIS

	Kapitel (erwähnt in Kapitel)
Abbildung 1: Messprinzip vom Ebio	3.1.1.1
Abbildung 2: Strom-Zeit-Kurve	3.1.1.1
Abbildung 3: X/Y-Diagramm bei linearer Regression	3.3.2
Abbildung 4: Normalisierter Differenzenplot	3.3.2
Abbildung 5: Prinzip des Diskordanzraten-Tests	3.3.3.1
Abbildung 6: Analytische Wahrscheinlichkeiten positiver Diskordanzen	3.3.3.2
Abbildung 7: Analytische Wahrscheinlichkeiten negativer Diskordanzen	3.3.3.2
Abbildung 8: Dichteschätzung + positive Diskordanzkurve,	3.3.3.2
Abbildung 9: Dichteschätzung + Diskordanzkurven, Beispiel	3.3.3.2
Abbildungen 10 a - d: Mittelwerte und Standardabweichungen der gemessenen Glukosekonzentrationen der Standard- und Kontroll-Lösungen	4.1.1
Abbildung 11: Präzision von-Tag- zu-Tag	4.1.1
Abbildungen 12 – 25: Regressionsanalysen + Differenzenplots	8.2 (4.2.2)
Abbildungen 26 a – f: Intrasystemische Vergleiche von Ebio. Graphische Darstellung der Ergebnisse des Diskordanzraten-Tests	4.2.3.1
Abbildungen 27 a - f: Diskordanzraten (%) der intrasystemischen Vergleiche	4.2.3.1
Abbildungen 28 - 41: Intersystemische Vergleiche. Graphische Darstellung der Ergebnisse des Diskordanzraten-Tests	8.2 (4.2.3.2)
Abbildungen 42 a – f: Diskordanzraten (%) der intersystemischen Vergleiche im Kapillarblut	4.2.3.2
Abbildungen 43 a – f: Diskordanzraten (%) der intersystemischen Vergleiche im venösen Vollblut	4.2.3.2
Abbildungen 44 a – f: Diskordanzraten (%) der intersystemischen Vergleiche im venösen Plasma	4.2.3.2

8.5 ABKÜRZUNGEN

0	Nüchternzeitpunkt
120	Zeitpunkt 2 Stunden nach Glukosebelastung
2h-postprandial	2 Stunden nach Glukosebelastung
µl	Mikroliter
BÄK	Bundesärztekammer
BMI	Body Mass Index
CB	Kapilläres Vollblut
CB0	Nüchternes Kapillarblut
CB120	Postprandiales Kapillarblut (zwei Stunden nach Glukosebelastung)
DR	Diskordanzrate
DR(ges)	Gesamt-Diskordanzrate
DR(neg)	negative Diskordanzrate
DR(pos)	positive Diskordanzrate
DR-Test	Diskordanzraten-Test
EG	Entscheidungsgrenze
EG im/md	Entscheidungsgrenze zwischen intermediärer und diabetischer Stoffwechsellage
EMV	Entscheidungsorientierter Methodenvergleich
f	feminin
fn	falsch negativ
fp	falsch positiv
g/l	Gramm pro Liter
Glu-Alb	proteinhaltiger (Albumin-haltiger) Glukose-Standard
Glu-NIST	Glukose-Standard, angesetzt mit Glukose des National Institute of Standards and Technology
H	Stunde(n)
im	intermediär
ISO	International Standard Organisation
korr	korrigiert
m	maskulin
md	mit Diabetes
md ₉₅ bzw. md(95)	95%-Dispersion der Residuen (mean deviation of 95%)
mg/dl	Milligramm pro Deziliter
ml	Milliliter
MW	Mittelwert(e)
MW-Abweichung	Mittelwert-Abweichung
n	Anzahl
NaCl	Natrium-Chlorid
NIST	National Institute of Statistics and Technology
oGTT	oraler Glukosetoleranztest

P/B Regression	Regression nach Passing und Bablok
p_a	"analytische" Wahrscheinlichkeit (probability)
POCT	Point of care testing
Pop-Koll	Populationskollektiv
postprandial	Zwei Stunden nach Glukosebelastung
p_p	Populationswahrscheinlichkeit
pTk	pro Tag korrigiert
pT-korrigiert	pro Tag korrigiert
RiliBÄK	Richtlinie(n) der Bundesärztekammer
roh	unkorrigiert
rn	richtig negativ
rp	richtig positiv
SRM	Standard Reference Material
Stud-Koll	Studienkollektiv
$s_{y,x}$	Standardabweichung der Residuen
trf	transformiert
uk	unkorrigiert
Val A	Validate A
Val N	Validate N
VB	Venöses Vollblut
VB0	Venöses Blut zum Nüchtern-Zeitpunkt
VB120	Venöses Blut zum Postprandial-Zeitpunkt (zwei Stunden nach Glukosebelastung)
VK%	Variationskoeffizient
VP	Venöses Plasma
VP0	Venöses Plasma zum Nüchtern-Zeitpunkt
VP120	Venöses Plasma zum Postprandial-Zeitpunkt (zwei Stunden nach Glukosebelastung)
vs	versus
WHO	World Health Organization
ZP	Zeitpunkt
ZP-0	Nüchternzeitpunkt
ZP-120	Zeitpunkt zwei Stunden nach Glukosebelastung (postprandial)

8.6 LEBENS LAUF

Persönliche Daten

Name: Püntmann
Vorname: Isabel Rosa Christiane
Geburtsdatum / -ort: 28.07.1970 in Fulda
Familienstand: ledig

Schulbildung

1977 – 1981 Grundschule Biebertal Hofbieber
1981 – 1990 Gymnasium Marienschule Fulda (Abschluss: Allgemeine Hochschulreife)

Hochschulbildung

1991 – 1998 Studium der Humanmedizin an der Georg-August Universität, Göttingen
1999 – 2001 Ärztin im Praktikum im Martins-Krankenhaus gGmbH, Lilienthal
(12 Monate Chirurgie und 6 Monate Innere Medizin)
März 2001 Vollapprobation

Berufliche Tätigkeiten

2001 – 2002 hausärztliche KV-Notdienste in Bremen (Raum Weyhe und Stuhr)
2002 – 2007 Assistenzärztin im Institut für Pharmakologie am Klinikum Bremen-Mitte gGmbH
2007 Assistenzärztin auf der Stroke Unit am Klinikum Bremen-Mitte gGmbH
2007 bis jetzt Assistenzärztin im Institut für Pharmakologie am Klinikum Bremen-Mitte gGmbH

Wissenschaftliche Publikationen

Haeckel R, Püntmann I. Validation of measurement procedures. In vitro veritas 2001; (2):
<http://www.acclc.cat/invitroveritas/vol2/art20.html>.

Püntmann I, Wosniok W, Haeckel R. Application of the diagnostic discordance rate in comparing various POCT glucometers for the diagnosis of type 2 diabetes. Clinical Chemistry and Laboratory Medicine 2001; (39): A60. Gemeinsame Jahrestagung der Deutschen Gesellschaft für Klinische Chemie (DGKC) und der Deutschen Gesellschaft für Laboratoriumsmedizin (DGL) 2001.

Püntmann I, Wosniok W, Haeckel R. A new test for the diagnostic relevance in method comparisons (diagnostic discordance rate). Clinical Chemistry and Laboratory Medicine 2001; (39): A60. Gemeinsame Jahrestagung der Deutschen Gesellschaft für Klinische Chemie (DGKC) und der Deutschen Gesellschaft für Laboratoriumsmedizin (DGL) 2001.

Püntmann I, Wosniok W, Haeckel R. Ein neues Maß für die diagnostische Relevanz beim Methodenvergleich. In: Medizinische Forschung und Gesundheitswissenschaften in Bremen. (Eds. Lison AE, Diehl HA), Universitätsdruckerei Bremen, 2001. 2. Symposium Medizinische Forschung und Gesundheitswissenschaften in Bremen 2001.

Püntmann I, Mühlbauer B. Rituximab, ein neues Therapieprinzip beim Non-Hodgkin-Lymphom. Arzneiverordnung in der Praxis 2002;4.

Püntmann I, von Maxen A, Mühlbauer B. Evidenzbasierte Medizin - Viel Leid mit Leitlinien. Bremer Ärztejournal 2002; (55): 17.

- Haeckel R, Brinck U, Colic D, Janka H U, Püntmann I, Schneider J, Viebrock C. Comparability of blood glucose concentrations measured in different sample systems for detecting glucose intolerance. *Clinical Chemistry* 2002; (48): 936-939. <http://www.clinchem.org/cgi/reprint/48/6/936>
- Haeckel R, Wosniok W, Puentmann I. Discordance rate, a new concept for combining diagnostic decisions with analytical performance characteristics. 1. Application in method or sample system comparisons and in defining decision limits. *Clinical Chemistry and Laboratory Medicine* 2003; (41): 347-355.
- Püntmann I, Wosniok W, Haeckel R. Comparison of several point-of-care testing (POCT) glucometers with an established laboratory procedure for the diagnosis of type 2 diabetes using the discordance rate. A new statistical approach. *Clinical Chemistry and Laboratory Medicine* 2003; (41): 809-820.
- Püntmann I, Wille H. COMET-Studie - Welchen Betablocker bei Herzinsuffizienz? *Bremer Ärztejournal* 2003; (56): 17.
- Püntmann I, Mühlbauer B. Arzneimittelausgabenbegrenzungsgesetz - "... oder inhaltsgleiche Präparate ..." - Konsequenzen der Entlassungsmedikation. *Bremer Ärztejournal* 2003; (56): 16.
- Boomgaarden-Brandes K, Langheit D, Püntmann I. Unerwünschte Arzneimittelwirkungen - Störwirkungen bedenken. *Bremer Ärztejournal* 2004; (57): 17.
- Püntmann I, Mühlbauer B, Wille H. Acute severe psychosis following testosterone substitution in klinefelter syndrome. *Naunyn-Schmiedeberg's Archives of Pharmacology* 2004; (369): R157 (625). 45. Frühjahrstagung der Deutschen Gesellschaft für Experimentelle und Klinische Pharmakologie und Toxikologie (DGPT) 2004.
- Mühlbauer B, Püntmann I, Wille H. Misteltoe-Induced Autoimmune-Hepatitis. *Naunyn-Schmiedeberg's Archives of Pharmacology* 2004; (369): R157 (628). 45. Frühjahrstagung der Deutschen Gesellschaft für Experimentelle und Klinische Pharmakologie und Toxikologie (DGPT) 2004.
- Boomgaarden-Brandes K, Püntmann I, Wille H. Was belegt PROVE-IT? *Bremer Ärztejournal* 2004; (57): 12.
- Püntmann I, Wolter S, Arzideh F, Wessels J, Albrecht K, Struff W G, Timm J, Mühlbauer B. Safety and Efficacy of a Hepatitis B Immunoglobulin Preparation Combined with Hepatitis B Vaccination in Neonates. Abstract-Band 9th Biennial Congress of the ESDPPP 2004;115 (P33). 9th Biennial Congress of the European Society for Developmental Perinatal + Paediatric Pharmacology (ESDPPP) 2004.
- Püntmann I, Wolther S, Arzideh F, Wessels J, Albrecht K, Struff W G, Timm J, Mühlbauer B. Efficacy and Safety of Hepatitis B Prophylaxis in Neonates Born to HBsAg Positive Mothers. *International Journal of clinical pharmacology and therapeutics* 2004; (42): 623. 6. Jahreskongress des Verbund Klinische Pharmakologie Deutschland (V Kli Pha) 2004.
- Langheit D, Püntmann I, Mühlbauer B. Systematic registration and documentation of drug consultations via a conceptual designed database. *International Journal of clinical pharmacology and therapeutics* 2004; (42): 617. 6. Jahreskongress des Verbund Klinische Pharmakologie Deutschland (V Kli Pha) 2004.
- Burmester G, Wille H, Püntmann I. Mystische Misteln. *Bremer Ärztejournal* 2005; (58): 13.
- Burmester G, Püntmann I, Boomgaarden-Brandes K, Langheit D, Wille H, Mühlbauer B. Evidenzbasierte Medizin in der Praxis. Evidenzbasierte Arzneimittelberatung (AMB) - Ein Brückenschlag zwischen Theorie und Praxis? - Vorstellung der Arzneimittelberatung in Bremen anhand Daten der Arzneimitteldatenbank. *Zeitschrift für ärztliche Fortbildung und Qualität im Gesundheitswesen* 2005; (100): 226 (13.12). 6. Jahrestagung des Deutschen Netzwerkes Evidenzbasierte Medizin (DNEbM) 2005.
- Püntmann I, Langheit D, Mühlbauer B. Clopidogrel vs. ASS + Esomeprazol - ASS-bedingtes Ulkus - Was tun? *Bremer Ärztejournal* 2005; (58): 22.
- Püntmann I, Breier-Wolski C, Mühlbauer B. Symposium in Bremen - Perspektiven der Pharmakovigilanz in Deutschland. *Bremer Ärztejournal* 2005; (58): 14.
- Püntmann I, Burmester G. Medizinische Adressen im Internet - Hilfe aus dem weltweiten Netz. *Bremer Ärztejournal* 2005; (58): 13.
- Jackisch-Riemann A, Püntmann I, Mühlbauer B. Pioglitazon und die PROactive-Studie - Weniger Antworten als Fragen. *Bremer Ärztejournal* 2005; (58): 14.
- Wille H, Püntmann I. Vitamine-Zusätze - Mehr Schaden als Nutzen. *Bremer Ärztejournal* 2006; (59): 14.

- Püntmann I, Wille H. Weitere Langzeitdaten nötig - Tiotropium bei COPD. Bremer Ärztejournal 2006; (59): 13.
- Püntmann I. Evidenzbasierte Patienteninformationen - Verbirgt sich hinter dem Etikett etwas Neues? Bremer Ärztejournal 2007; (60): 15-16.
- Püntmann I, Bobis Seidenschwanz I, Wille H, Mühlbauer B. Panic attacks, depression and suicidal ideation for 5 months after only one dose of prophylactic mefloquine. Naunyn-Schmiedeberg's Archives of Pharmacology 2007; (372). 48. Frühjahrstagung der Deutschen Gesellschaft für Experimentelle und Klinische Pharmakologie und Toxikologie (DGPT) 2007.
- Jackisch-Riemann A, Püntmann I, Bobis Seidenschwanz I, Wille H, Mühlbauer B. Osteonecrosis of the jaws in two patients treated with Zoledronic Acid (ZA). Naunyn-Schmiedeberg's Archives of Pharmacology 2007; (375). 48. Frühjahrstagung der Deutschen Gesellschaft für Experimentelle und Klinische Pharmakologie und Toxikologie (DGPT) 2007.
- Püntmann I, Wille H. Raucherentwöhnung - Vareniclin: Partieller Nikotin-Rezeptoragonist. Bremer Ärztejournal 2007; (60): 17.
- Ahrens W, Hagemeyer C, Mühlbauer B, Pigeot I, Püntmann I, Reineke A, Steinbach M, Timm J. Hospitalization rates of generic metoprolol compared with the original beta-blocker in an epidemiological database study. Pharmacoepidemiology and Drug Safety 2007; (16): 1298-1307.
- Püntmann I, von Maxen A, Wille H. Amlodipin - Hinweise zu Indikation und therapeutischem Nutzen. AOK Bundesverband. AOK-Bundesverband für Information 2 (Überarbeitung der Version von 2001), 1-6. 2008. http://www.aok-beratungsapotheke.de/05_tippstricks/documents/Amlodipin-14_5.pdf
- Püntmann I, Wille H. Zwei Antidiabetika mit neuartigem Wirkprinzip - Bisher ohne erkennbaren therapeutischen Stellenwert. Bremer Ärztejournal 2008; (61): 17.
- Püntmann I. Biologika bei Psoriasis - Alarmierende Nebenwirkungen. Bremer Ärztejournal 2008; (61): 12.
- (Die Bremer Ärztejournal-Artikel sind auf <http://www.pharmakologie-bremen.de> unter Publikationen hinterlegt)

Bremen, den

Isabel Püntmann

Ergebnisse der Dissertation wurden in folgenden Publikationsorganen veröffentlicht:

- (1) Haeckel R, Püntmann I. Validation of measurement procedures. *In vitro veritas* 2001; 2(art 20): <http://www.acclc.cat/invitroveritas/vol2/art20.html>.
- (2) Püntmann I, Wosniok W, Haeckel R. Application of the diagnostic discordance rate in comparing various POCT glucometers for the diagnosis of type 2 diabetes. *Clinical Chemistry and Laboratory Medicine* 2001; 39(9):A60.
Gemeinsame Jahrestagung der Deutschen Gesellschaft für Klinische Chemie (DGKC) und der Deutschen Gesellschaft für Laboratoriumsmedizin (DGL), 2001.
- (3) Püntmann I, Wosniok W, Haeckel R. A new test for the diagnostic relevance in method comparisons (diagnostic discordance rate). *Clinical Chemistry and Laboratory Medicine* 2001; 39(9):A60.
Gemeinsame Jahrestagung der Deutschen Gesellschaft für Klinische Chemie (DGKC) und der Deutschen Gesellschaft für Laboratoriumsmedizin (DGL), 2001.
- (4) Püntmann I, Wosniok W, Haeckel R. Ein neues Maß für die diagnostische Relevanz beim Methodenvergleich. In: Lison AE, Diehl HA, editors. *Medizinische Forschung und Gesundheitswissenschaften in Bremen*. Universitätsdruckerei Bremen, 2001.
2. Symposium Medizinische Forschung und Gesundheitswissenschaften in Bremen, 2001.
- (5) Haeckel R, Wosniok W, Puentmann I. Discordance rate, a new concept for combining diagnostic decisions with analytical performance characteristics. 1. Application in method or sample system comparisons and in defining decision limits. *Clinical Chemistry and Laboratory Medicine* 2003; 41(3):347-355.
- (6) Püntmann I, Wosniok W, Haeckel R. Comparison of several point-of-care testing (POCT) glucometers with an established laboratory procedure for the diagnosis of type 2 diabetes using the discordance rate. A new statistical approach. *Clinical Chemistry and Laboratory Medicine* 2003; 41(6):809-820.

Bremen, den

(Isabel Püntmann)

9 DANKSAGUNG

Mein ganz besonderer Dank gilt

- Herrn Prof. Dr. med. Rainer Haeckel für die Überlassung der Arbeit, die engagierte Betreuung während des praktischen Teils meiner Arbeit, die schnelle Durchsicht und Korrektur meiner Entwürfe und die immer wiederkehrende Bereitschaft, sich mit dem Thema zu befassen. Durch ihn habe ich Einblick in wissenschaftliches Arbeiten und Publizieren erhalten.

- den Mitarbeiterinnen und Mitarbeitern seines Instituts, die immer hilfreich zur Seite standen und eine freundliche Arbeitsatmosphäre geschaffen haben.

- dem an der Studie beteiligten Biometriker Herrn Werner Wosniok, der mir mit verständlichen Worten den Diskordanzratentest erklärte und kontinuierlich an der vereinfachten Anwendung des zunächst in SAS entwickelten und dann auf Excelbasis umgewandelten Rechenmoduls gearbeitet hat.

- Herrn Prof. Dr. med. Bernd Mühlbauer für hilfreiche Anregungen und Ratschläge und Herrn Dr. med. Wolfgang Flocken für die Durchsicht und Korrektur des Manuskripts.

- Herrn Dr. med. Eberhard Otto, meinen Kollegen und meinen Freunden, insbesondere Elke Heussler, Stine Voß, Erwin Ernst Starke, Rüdiger Otto, Winfried Doehring und Dr. med. Anne Jackisch-Riemann für ihre Geduld, ermunternden Worte und fortwährende Unterstützung.

- den Patienten für ihre Einwilligung, an der Studie teilzunehmen, und ihre Mitarbeit.

- meinen Eltern, die mich unterstützt und an mich geglaubt haben. Meinem Dank ihnen gegenüber möchte ich Ausdruck verleihen, indem ich ihnen diese Arbeit widme.