

Aus dem Institut für Klinische Chemie und Laboratoriumsmedizin
(Direktor: Univ.- Prof. Dr. med. M. Nauck)
der Universitätsmedizin der Ernst-Moritz-Arndt-Universität Greifswald

**Einfluss der Messunsicherheit auf die Diagnose Diabetes mellitus:
Verlässliche Minimal Difference von Plasma Glukose Messungen.**

Inaugural – Dissertation
zur
Erlangung des akademischen Grades
Doktor der Medizin
(Dr. med.)

der Universitätsmedizin
der Ernst-Moritz-Arndt-Universität

Greifswald

2020

vorgelegt von:
Sandra Keutmann
geb. am: 12.11.1975
in: Köln

Dekan: Herr Prof. Dr. Karlhans Endlich
1. Gutachter: Frau Prof. Dr. Astrid Petersmann
2. Gutachter: Herr Prof. Dr. Erwin Schleicher
Ort, Raum: Zoom, Meeting- ID: 619 9056 3938
Tag der Disputation: 05.03.2021

Für meine Mutter

Inhalt

Abkürzungsverzeichnis.....	1
1. Einleitung	3
1.1 Messunsicherheit	4
1.2 Beschreibung der Minimal Difference.....	5
1.3 Zielsetzung	8
2. Material und Methoden.....	9
2.1 Minimal Difference.....	11
2.2 Langzeitauswertung	12
2.3 Kurzeitauswertung.....	12
2.4 Berechnungen und Statistiken.....	13
3. Ergebnisse	15
3.1 Langzeitauswertung	15
3.2 Kurzeitauswertung.....	18
4. Diskussion.....	22
4.1 Einschränkungen	26
4.2. Ausblick.....	27
6. Zusammenfassung.....	28
7. Literaturverzeichnis	29
8. Abbildungsverzeichnis.....	31
9. Tabellenverzeichnis.....	32
10. Formelverzeichnis.....	33
11. Wissenschaftliche Publikation.....	34
14. Danksagung	46

Abkürzungsverzeichnis

ADA	American Diabetes Association (Amerikanische Diabetesgesellschaft)
CI	Confidence Interval (Konfidenzintervall)
CLSI	Clinical and Laboratory Standards Institute
DDG	Deutsche Diabetes Gesellschaft
DIN EN ISO	Deutsche Industrienorm, Europäische Norm, International Organisation for Standardisation
DM	Diabetes mellitus
FPG	Fasting Plasma Glucose (Nüchternplasmaglukose)
GUM	Guide to the Expression of Uncertainty in Measurement
IFG	Impaired Fasting Glucose (gestörte Nüchternglukose)
IGT	Impaired Glucose Tolerance (gestörte Glukosetoleranz)
KHK	Koronare Herzkrankheit
LIS	Laborinformationssystem
QK	Interne Qualitätskontrolle
M	Mittelwert
MD	Minimal Difference
MIO	Millionen
pAVK	Periphere arterielle Verschlusskrankheit
OGTT	Oraler Glukosetoleranztest
Rili-BÄK	Richtlinie der Bundesärztekammer zur Qualitätssicherung laboratoriumsmedizinischer Untersuchungen
SD	Standardabweichung

VK Variationskoeffizient

WHO World Health Organisation (Weltgesundheitsorganisation)

1. Einleitung

Diabetes mellitus (DM) ist eine Gruppe heterogener Erkrankungen des Stoffwechsels, deren gemeinsames Merkmal die chronische Hyperglykämie ist. Die Ursache ist entweder eine gestörte Insulinsekretion des Pankreas oder eine gestörte Insulinwirkung der Zielgewebe oder beides (1-3). DM ist eine Erkrankung, die in den letzten Jahren kontinuierlich zugenommen hat. Laut der Weltgesundheitsorganisation (WHO) betraf es im Jahr 2005 weltweit 171 Millionen (Mio.) Menschen, davon über 7 Mio. in Deutschland. Es wird angenommen, dass sich diese Zahl bis zum Jahr 2030 auf über 300 Mio. Menschen verdoppeln wird (4-6). Etwa 5-10% der Patienten mit einem DM leiden an einem Typ-1-DM während ca. 90% an einem Typ-2-DM leiden (7). Während früher an dem Typ-2-DM überwiegend ältere Menschen erkrankt sind, bekommen heute zum Teil sehr junge Menschen diese Diagnose. Als Risikofaktoren gelten vor allem Bewegungsmangel und Übergewicht (8). Die Krankheit hat schwerwiegende Komplikationen, unter denen Millionen von Menschen leiden (9). Die Langzeitschäden sind mit Erkrankungen von verschiedenen Organen assoziiert, wie z.B. die Retinopathie, die Nephropathie mit terminalem Nierenversagen, die periphere Neuropathie, sowie die koronare Herzkrankheit (KHK), die periphere Verschlusskrankheit (pAVK), gastrointestinale und urogenitale Organschäden (10-12). Die Folgeerkrankungen sind für den Einzelnen ebenso aber für unser Gesundheitssystem eine Herausforderung, denn auch hier fallen erhebliche Kosten an (13). Schätzungen zufolge wird DM Typ 2 häufig erst 4-7 Jahre nach Krankheitsbeginn diagnostiziert; zu diesem Zeitpunkt haben aber bereits Gefäßveränderungen ihren Anfang genommen (14). Die Zahlen machen deutlich, wie wichtig Diagnosestellung und Therapie sind.

Die Diagnostik des DM erfolgt mit Hilfe geeigneter Biomarker anhand diagnostischer Entscheidungsgrenzen für die jeweilige Messgröße, z.B. für die Plasmaglukosekonzentration. Die Weltgesundheitsorganisation (WHO), die American Diabetes Association (ADA) sowie die Deutsche Diabetes Gesellschaft (DDG) empfehlen für die Bestimmung von Glukose venöse Plasmaproben.

Die diagnostische Entscheidungsgrenze für Nüchternplasmaglukose (FPG) beträgt ≥ 7.0 mmol/L (≥ 126 mg/dL) (15) und im Rahmen eines oralen Glukosetoleranztestes (OGTT) zwei Stunden nach oraler Aufnahme von 75 g Glukoselösung ≥ 11.1 mmol/L (≥ 200 mg/dL) (15-17). Eine gestörte Nüchternglukose (IFG) wird gemäß WHO im Intervall zwischen 6.1 mmol/L und 6.9 mmol/L (110 mg/dL bis 125 mg/dL) und gemäß ADA und DDG zwischen 5.6 mmol/L und 6.9 mmol/L (100 mg/dL bis 125 mg/dL) (15) Plasmaglukose diagnostiziert.

1.1 Messunsicherheit

Laborergebnisse werden oft wie ein Kontoauszug betrachtet: Der berichtete Wert gilt als wahr und bereits ein paar Cent mehr oder weniger würden nicht akzeptiert. Alle Messungen in der Laboratoriumsmedizin hingegen sind immer mit einer Messunsicherheit behaftet. Ein paar "Cent" mehr oder weniger in den Messungen können das gleiche Ergebnis darstellen, abhängig vom Ausmaß der assoziierten Messunsicherheit. Alle laboratoriumsmedizinischen Untersuchungen, einschließlich der Glukosekonzentrationsbestimmung, unterliegen analytischen Schwankungen und ein Messergebnis ist nur vollständig unter der Angabe der damit verbundenen Messunsicherheit. Dies wird ausführlich in den Standardwerken, wie „The Guide to the expression of uncertainty in measurements (GUM)“ beschrieben (18).

Trotzdem wird für laboratoriumsmedizinische Untersuchungen die Messunsicherheit i.d.R. nicht auf dem Befund mitgeteilt, sondern kann im Labor erfragt werden. Dabei wird in medizinischen Laboratorien die Impräzision im Allgemeinen als Variationskoeffizient (%VK) (19) angegeben. Der in Prozent angegebene VK (Gleichung 1) ist nur schwer in eine absolute Differenz zu übersetzen, die erforderlich wäre, um zwei aufeinanderfolgenden Ergebnissen als verschieden bezeichnen zu können. Ebenso kann der VK nicht unmittelbar bei der Entscheidung herangezogen werden, ob sich ein Messergebnis, vor dem Hintergrund der Messunsicherheit tatsächlich von einem Grenzwert unterscheidet.

Für diesen Zweck ist die Angabe der Impräzision in den Einheiten der Messgröße besser geeignet, z.B. in Form der Standardabweichung oder der "Minimal Difference". Es ist wichtig, dass der Kliniker die Impräzision berücksichtigt, wenn er ein Laborergebnis anhand einer diagnostischen Entscheidungsgrenze bewertet und für die Diagnosestellung des DM heranzieht. Für diesen Zweck ist die Minimal Difference besonders geeignet, da sie die Impräzision in der Einheit der zu beurteilenden Messgröße angibt.

1.2 Beschreibung der Minimal Difference

Die DDG hat erstmals 2017 in ihrer jährlich überarbeiteten Praxisempfehlung zur Definition, Klassifizierung und Diagnose des DM die Mindestanforderungen der erlaubten Messunsicherheit als Minimal Difference (MD) angegeben (17). Dabei handelt es sich um kein neues, sondern um ein etabliertes Konzept, das z.B. bereits in den Leitlinien des Clinical and Laboratory Standard Institute (CLSI) EP29-A zur „Expression of Measurement Uncertainty in Laboratory Medicine“ beschrieben wurde (20). Die MD steht für die kleinste Differenz zwischen zwei Werten, die notwendig ist, um sie als statistisch signifikant verschieden bezeichnen zu können, wenn man die zugrunde liegende Messunsicherheit berücksichtigt. Die MD basiert auf der Standardabweichung (SD), wird in der Einheit der Messgröße angegeben und kann direkt auf das Messergebnis angewendet werden (21).

Die DDG empfiehlt für die Diagnose des Diabetes mellitus bei der FPG-Grenze von 7.0 mmol/L eine $MD_{\text{cut-off}}$ von ≤ 0.7 mmol/L ($MD_{\text{cut-off } 7.0}$). Bei der Glukosekonzentration von 7.0 mmol/L entsprechen eine SD von 0.35 mmol/L und eine MD von 0.7 mmol/L (Gleichung 5) einem VK von $\leq 5\%$. Ein VK von 6% würde einer MD von 0.84 mmol/L entsprechen. Übertragen auf das Messergebnis könnte bei dieser Impräzision der wahre Wert zwischen 6.16 mmol/L und 7.84 mmol/L betragen. Angewendet auf eine Konzentration von 5.6 mmol/L würde das Intervall von 4.93 mmol/L bis 6.27 mmol/L reichen. Ein analytischer VK dieser Größenordnung oder darüber würde es daher nicht ermöglichen, zwischen Diabetes und Nicht-Diabetes zu unterscheiden. Dies wird durch die Überschneidung und Überlappung der Trichter in Abb.1 veranschaulicht.

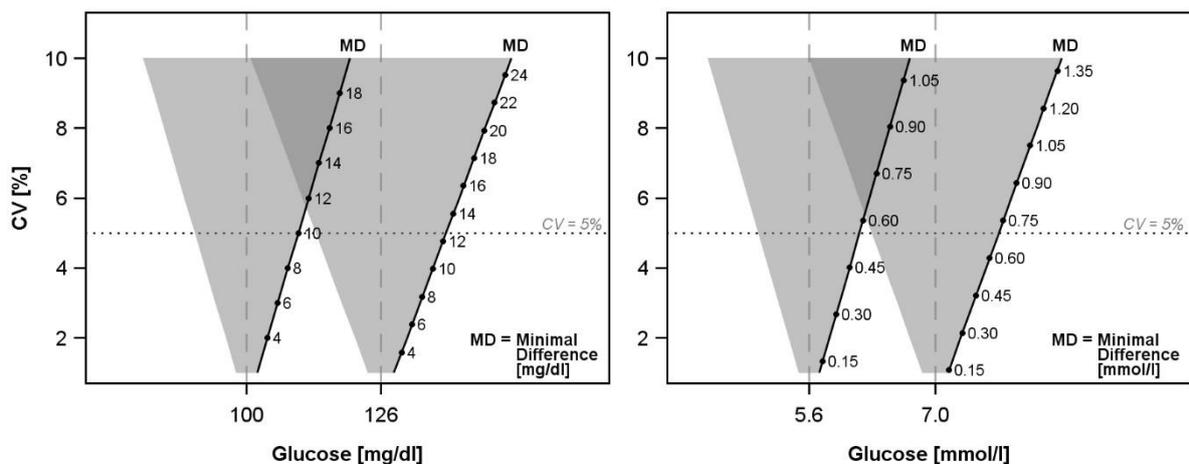


Abbildung 1: Beziehung zwischen VK und Minimal Difference (MD) an den diagnostischen Cut-offs der Glukosekonzentration, die zur Diagnose von DM empfohlen werden.

Die Breite der Trichter wird durch die Impräzision determiniert und veranschaulicht die Messunsicherheit in der Einheit der Messgröße Glukose. Die MD kann für die angegebenen Glukosekonzentrationen (x-Achse) an der Längsseite der Trichter für jeden VK (y-Achse) abgelesen werden. Die gepunkteten Linien markieren den maximal zulässigen VK von 5%, wie er von der DDG empfohlen wird (1). Die Überlappungsbereiche oberhalb der Schnittpunkte der Trichter entsprechen VKs, bei denen zwischen den Glukosekonzentrationen 5.6 und 7.0 mmol/L aufgrund der vorliegenden Impräzision nicht mehr unterschieden werden kann (1).

Qualitätssicherungsverfahren, wie die weit verbreiteten Westgard-Regeln (22) oder die in Deutschland rechtlich verbindliche Richtlinie der Bundesärztekammer zur Qualitätssicherung laboratoriumsmedizinischer Untersuchungen (Rili-BÄK) (23), erfordern mindestens zwei interne Qualitätskontrollen (QK) unterschiedlicher Konzentrationen pro Messgröße, Tag und Gerät. Diese Frequenz der QK setzt voraus, dass die Instrumente mit zuverlässig hoher Qualität und Stabilität betrieben werden können.

Wenn die Anforderungen des Kontrollsystems u.a. hinsichtlich der QK erfüllt sind, werden die Ergebnisse berichtet. Es wird angenommen, dass die Ergebnisse zwischen zwei akzeptierten Kontrollprobeneinzelmessungen den Qualitätskriterien entsprechen (19).

Darüber hinaus wird im Rahmen der Rili-BÄK die Messunsicherheit retrospektiv anhand sogenannter "Kontrollzyklen" (z.B. einem Monat) überwacht und bewertet (23).

Viele Labore nutzen mehrere Instrumente mit ähnlichem analytischem Spektrum parallel, um die hohe Anzahl von Patientenproben innerhalb einer akzeptablen Turnaround-Zeit zu messen oder um eine Sicherung für Ausfälle einzelner Instrumente bereitzuhalten. I.d.R. hat daher der Anforderer der laboratoriumsmedizinischen Untersuchung keine Kenntnis darüber, auf welchem Instrument die Messung durchgeführt wird.

1.3 Zielsetzung

Die vorliegende Studie untersucht die Messunsicherheit von Plasmaglukosekonzentrationsmessungen an einem kontinuierlich betriebenen Messsystem in einem Krankenhaus der Maximalversorgung. Die Glukosemessungen wurden an drei Instrumenten, die an eine Laborstraße angeschlossen sind, durchgeführt. Für die Studie wurden zwei Ansätze verfolgt:

1. Langzeitauswertung von QK-Daten zwischen 2009-2018 und
2. Kurzzeitauswertung mittels intensiviertem Schema, mit stündlichen Messungen von QK-Material über 7 Tage aufeinander folgend im September 2012.

Die MD der für die Diagnose von DM relevanten Glukosekonzentration wurde in beiden Ansätzen berechnet und verglichen.

Ziel der Studie war es, verlässliche Impräzisionsdaten für die Glukosekonzentrationsmessungen zu ermitteln und diese in die klinisch verwendbare MD zu übersetzen. Das Ausmaß der Messunsicherheit, mit der jedes Messergebnis unabdingbar behaftet ist, sollte für den Kliniker leicht ersichtlich sein, somit kann die Messunsicherheit bei der Diagnosestellung berücksichtigt werden, insbesondere, wenn Messwerte anhand fester diagnostischer Grenzwerte beurteilt werden, wie es bei DM der Fall ist.

2. Material und Methoden

Alle Messungen der vorliegenden Arbeit wurden am Institut für Klinische Chemie und Laboratoriumsmedizin der Universitätsmedizin Greifswald durchgeführt, das rund um die Uhr an sieben Tagen pro Woche laboratoriumsmedizinische Untersuchungen anbietet. Die Kurzzeitauswertung fand im September 2012 an sieben aufeinander folgenden Tagen im 24 h Rhythmus statt. Die Langzeitauswertung erfolgte für den Zeitraum Januar 2009 bis Dezember 2018. In diesem Zeitraum wurden drei Messsysteme (Dimension Vista) (Abb.2), verbunden durch eine Laborstraße (alle Siemens Healthcare GmbH, Eschborn, Deutschland), für die Patientenversorgung verwendet. Die Laborstraße wurde 2012 von StreamLab auf FlexLab umgestellt (Instrument 4; Siemens Healthcare GmbH, Eschborn, Deutschland). Beide Laborautomatisierungssysteme verteilen die Proben an die analytischen Dimension Vista-Instrumente 1 bis 3 entsprechend der Verfügbarkeit von Analysekapazitäten.

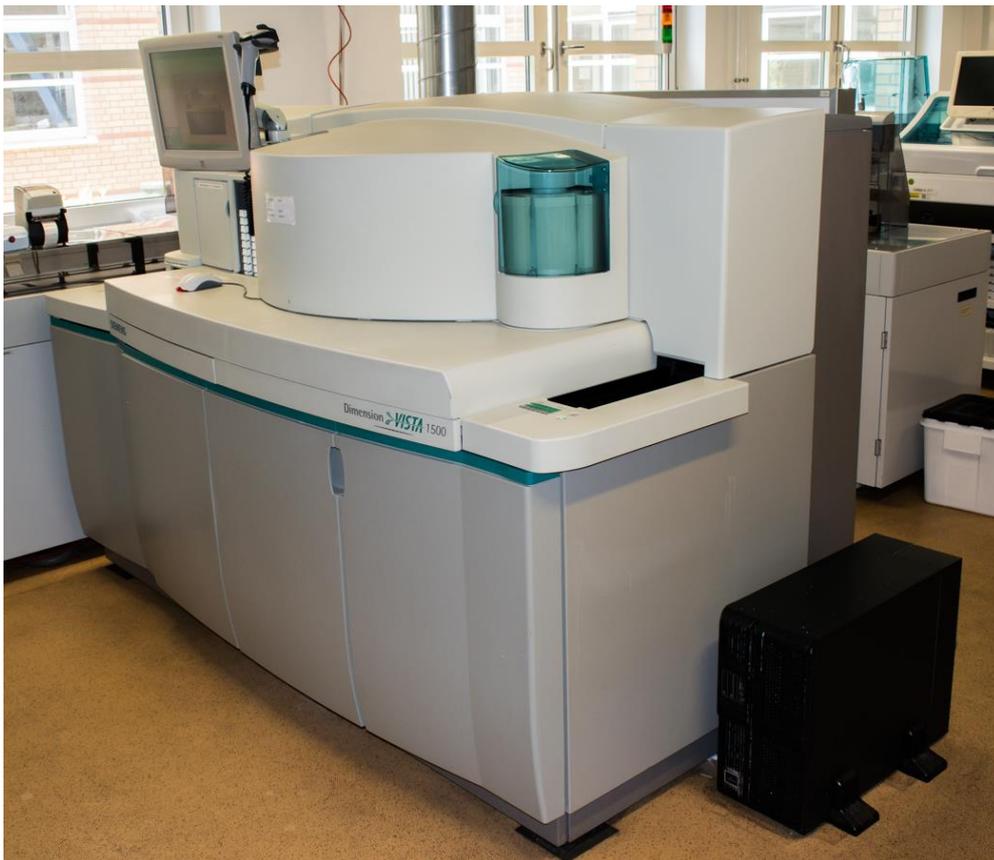


Abbildung 2: Dimension Vista angebunden an die Laborstraße FlexLab (Siemens Healthcare GmbH, Eschborn, Deutschland) im Institut für Klinische Chemie und Laboratoriumsmedizin der Universitätsmedizin Greifswald.

Die Plasmaglukosekonzentration wurde mit der Hexokinase-Methode (24), ebenfalls Siemens Healthcare GmbH, unter Einhaltung der Rili-BÄK bestimmt.

Die Hexokinase-Methode ist eine allgemein anerkannte spezifische Labormethode zur Messung von Glukose. Sie wird u.a. zur Messung der Plasmaglukosekonzentration in der Diagnose und Behandlung von Diabetes mellitus eingesetzt.

Das Enzym Hexokinase setzt Hexosen und Adenosintriphosphat (ATP) und Magnesium zu Hexose-6-phosphat und Adenosindiphosphat (ADP) um. Anschließend dehydriert das Nicotinamid-adenin-dinucleotid-phosphat (NADP⁺) in Gegenwart des Enzyms Glukose-6-Phosphat-dehydrogenase (G6P-DH), das entstandene Glukose-6-phosphat zu D-Glukonat-6-phosphat. Dabei wird das reduzierte Nikotinamid-adenin-dinucleotid-phosphat (NADPH) gebildet. Wie in der Reaktionsgleichung (Abbildung 3) zu erkennen, ergibt eine Ausgangsmenge an D-Glukose eine äquivalente Stoffmenge NADPH. NADPH kann photometrisch bestimmt werden, da NADPH im Gegensatz zu NADP⁺ bei 340 nm ein Absorptionsmaximum besitzt.

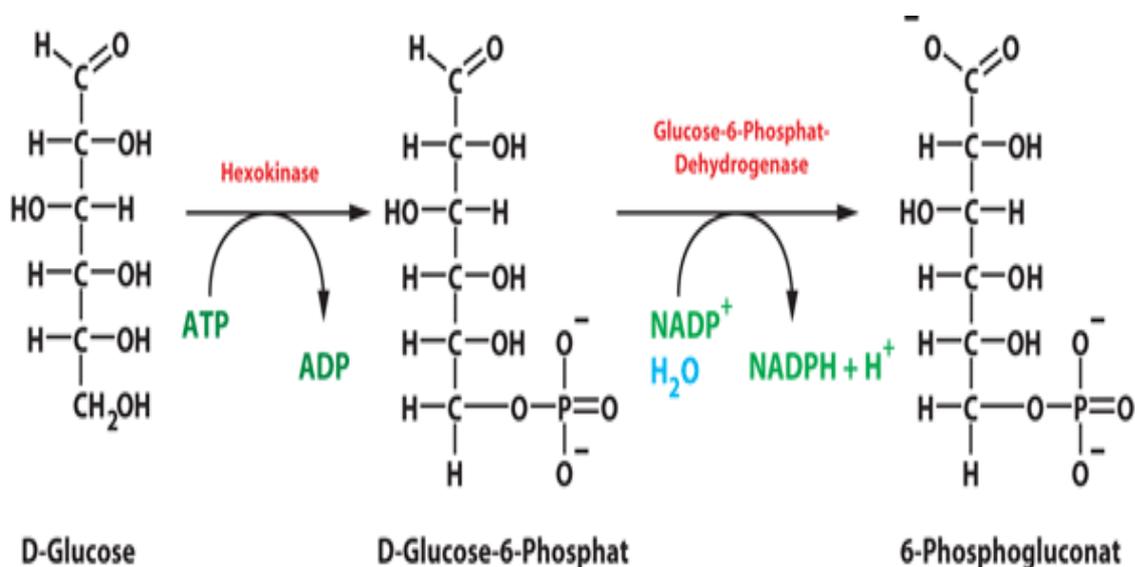


Abbildung 3: Hexokinase-Glukose-6-Phosphat-Dehydrogenase-Methode

Für diese Studie wurde ausschließlich kommerziell erhältliches Qualitätskontrollmaterial und damit kein Patientenmaterial verwendet. Daher war kein Ethikvotum erforderlich. Von Januar 2009 bis Dezember 2013 wurde das QK-Material Tru Liquid Monitrol (Thermo Fisher Scientific, Schwerte, Deutschland) verwendet. Ab Januar 2014 bis Dezember 2018 wurde das Kontrollmaterial Liquid Assayed Multiqua (BIO-RAD Laboratories, München, Deutschland) eingesetzt.

2.1 Minimal Difference

Die MD basiert auf der SD und wird in einer Einheit der betreffenden Messgröße ausgedrückt, so dass Kliniker die MD direkt auf das Messergebnis anwenden können (21). Wird für die Berechnung der MD $k = 2$ verwendet, so entspricht dies einem Konfidenzniveau von 95% (26). Wenn zwei aufeinander folgende Messungen miteinander verglichen werden, muss die Messunsicherheit beider Messungen berücksichtigt werden (Gleichung 2). Wird ein Messergebnis mit einer festgelegten diagnostischen Entscheidungsgrenze verglichen (Gleichung 3), so muss nur die Messunsicherheit des Messergebnisses berücksichtigt werden, da die festgelegte Entscheidungsgrenze von Natur aus nicht mit einer Messunsicherheit behaftet ist. Wird also ein Messergebnis mit einem festen Grenzwert verglichen (Gleichung 4), so kann die MD zu MD-Grenzwert ($MD_{cut-off}$) vereinfacht werden (Gleichung 5).

$$\%VK = \frac{SD}{M} * 100; SD = \frac{M * \%VK}{100} \quad (1)$$

$$MD = k \times \sqrt{SD_1^2 + SD_2^2} \dots\dots\dots (2)$$

wenn $SD_1 = SD_2$ (Annahme bei zwei aufeinander folgenden Messwerten)

$$MD = k \times SD \times \sqrt{2} \dots\dots\dots (3)$$

wenn $SD_2 = 0$ (diagnostischer Cut-off)

$$MD_{cut-off} = k \times \sqrt{SD^2} \dots\dots\dots (4)$$

$$MD_{cut-off} = k \times SD = 2 \times SD \quad (5)$$

Formel 1: Gleichung (1)-(5); $k=2$ entspricht einem Konfidenzniveau von 95%

2.2 Langzeitauswertung

Impräzision

Dokumentierte QK-Daten aus dem Laborinformationssystem (LIS) wurden retrospektiv für den Studienzeitraum von zehn Jahren ausgewertet. Die Kontrollzyklen wurden monatlich für jedes Instrument und QK-Level separat ausgewertet. Zyklen mit weniger als 15 QK-Ergebnissen wurden ausgeschlossen. Der $MD_{\text{cut-off}}$ wurde monatlich nach der Formel 1 (Gleichungen 1-5) berechnet.

2.3 Kurzzeitauswertung

Arbeitslast

Die Gesamtzahl aller möglichen Messungen der internen QK gemäß des Studiendesigns betrug 2.016 über alle Instrumente. Die tatsächliche Anzahl von Messergebnissen wurde dokumentiert. Eine Reduktion der Messungen konnte z.B. aufgrund von Wartungen auftreten, die täglich erforderlich waren. Die Gesamtarbeitslast, d.h. die Anzahl aller Analysen, wurde rückwirkend für jedes System aus dem LIS abgerufen.

Impräzision

Separates Kontrollmaterial wurde im September 2012 über einen Zeitraum von einer Woche stündlich gemessen. Die Proben wurden direkt in jedes Gerät 1-3 oder zufällig durch FlexLab (Instrument 4) eingeführt. Die Eingabe über die Laborstraße FlexLab soll den allgemeinen Prozessablauf für Patientenproben repräsentieren, da diese i.d.R. nicht direkt manuell in die Geräte gestellt werden. Die Qualitätssicherung und Qualitätsüberwachung erfolgte unabhängig vom Kontrollmaterial der Kurzzeitauswertung mit separatem internen QK-Material gemäß Rili-BÄK. Es wurden nur Messungen eingeschlossen, die die Rili-BÄK erfüllten. Die Messungen des separaten Kontrollmaterials waren demselben Probenfluss unterworfen wie die Patientenproben und die Ergebnisse wurden nach Studienende aus dem Laborinformationssystem (LIS) extrahiert. Um Kontrollzyklen für die Kurzzeitstudie zu bilden, wurden die stündlichen Messungen eines Tages zusammengefasst, separat für jedes Instrument und Konzentrationslevel.

Es wurden drei Konzentrationslevel von etwa 3, 7 und 20 mmol/L (BIO-RAD Liquid Assayed LOT 45631-1 bis 3, BIO-RAD Laboratories München, Deutschland) verwendet. Für jedes Level wurde ausreichend Kontrollmaterial aufgetaut und gepoolt, um den Bedarf für einen Tag zu decken. Anschließend erfolgte die Aliquotierung in barcodierte Röhrchen, die zur Verwendung an den Instrumenten geeignet waren. Bis zur Messung wurde das für einen Studientag aliquotierte Material verschlossen bei 2-5°C gelagert, maximal jedoch 24 Stunden.

2.4 Berechnungen und Statistiken

Die Berechnung von VK und MD einschließlich deskriptiver Statistiken und Boxplots wurden mit R Studio (Version 3.5.0; 2018-04-23, Release 23. März 2018; <https://www.r-project.org/>) durchgeführt. Statistische Unterschiede zwischen den Instrumenten und Jahren wurden unter Verwendung der Varianzanalyse berechnet und Tukeys HSD-Test wurde als Post-Hoc-Test verwendet. Die Auslastungsgrafiken wurden mit der SAS-Software Version 9.4 (SAS Institute Inc, Cary, North Carolina, USA) durchgeführt. Die MD wurde mit der Formel 1 (Gleichung (5) unter der Annahme erstellt, dass der Cut-off keiner Impräzision unterliegt.

In der Langzeitauswertung wurden die MDs monatlich für jedes Instrument unter Verwendung der QK-Daten berechnet und dann als Moving Average kumuliert, um einen verlässlichen Wert zu erhalten (25). In der Kurzzeitauswertung wurden die MDs täglich für jedes Instrument berechnet.

Die Gleichungen (1) und (5) wurden für beide Ansätze verwendet. Die Ergebnisse wurden in Boxplots für beide Studienansätze zusammengefasst und nach Jahren bzw. Instrumenten detailliert dargestellt.

Um zuverlässige Werte für die Minimal Difference zu erhalten, die für die klinische Verwendung bei der Diagnose von Diabetes mellitus geeignet ist, wurden die 95%-Perzentile der $MD_{\text{cut-off}}$ Verteilungen verwendet. Der $MD_{\text{cut-off}}$ der verschiedenen in dieser Studie verwendeten Glukosekonzentrationen wurde herangezogen, um eine lineare Gleichung zu berechnen, mit der der $MD_{\text{cut-off}}$ für andere Konzentrationen, wie der $MD_{\text{cut-off}}$ bei 7.0 mmol/L, abgeleitet werden kann.

Da das im Kurzzeitexperiment verwendete Material gepoolt und dann auf die Instrumente verteilt wurde, war eine Berechnung der systematischen Abweichung zwischen den Instrumenten möglich. Diese Abweichung wurde für jedes Instrument separat berechnet, indem die mittlere $MD_{\text{cut-off}}$ von jedem Instrument von der medianen $MD_{\text{cut-off}}$ aller Instrumente abgezogen wurde.

3. Ergebnisse

3.1 Langzeitauswertung

Grundlage für die Langzeitauswertung der Messunsicherheit bei der Bestimmung der Glukosekonzentration waren die Qualitätskontrollmessungen, die im Rahmen der Patientenversorgung gemäß Rili-BÄK durchgeführt wurden. Die ermittelten Werte für die MD konvergierten und erreichten nach 30 Beobachtungen ein stabiles Niveau. Alle Ergebnisse wurden Rili-BÄK konform gewonnen und in Abbildung 4. zusammengefasst.

Die Größe des $MD_{\text{cut-off}}$ hängt von der SD ab und nimmt mit steigenden Konzentrationen zu, wie z.B. an der linearen Regression in Abbildung 4 zu sehen. Der $MD_{\text{cut-off } 7.0}$ wurde aus der linearen Gleichung $y = 0,065 x - 0.013$ mmol/L anhand von vier Konzentrationen (~3.4, ~5.5, ~14.7, und ~19.1 mmol/L) abgeleitet und mit 0.44 mmol/L berechnet (Abbildung 4c). Die $MD_{\text{cut-off } 7.0}$ unterschied sich geringfügig, aber signifikant zwischen den Instrumenten (Abbildung 5).

Ein geringfügiger Unterschied wurde für einige QKs gefunden, die im Jahr 2009 höher waren als in allen anderen Jahren. Nach der Umstellung des QK-Materials im Jahr 2014 war die Performance über alle untersuchten Jahre stabil, d.h. es wurden keine signifikanten Unterschiede in den $MD_{\text{cut-off } 7.0}$ der einzelnen Jahre mehr beobachtet.

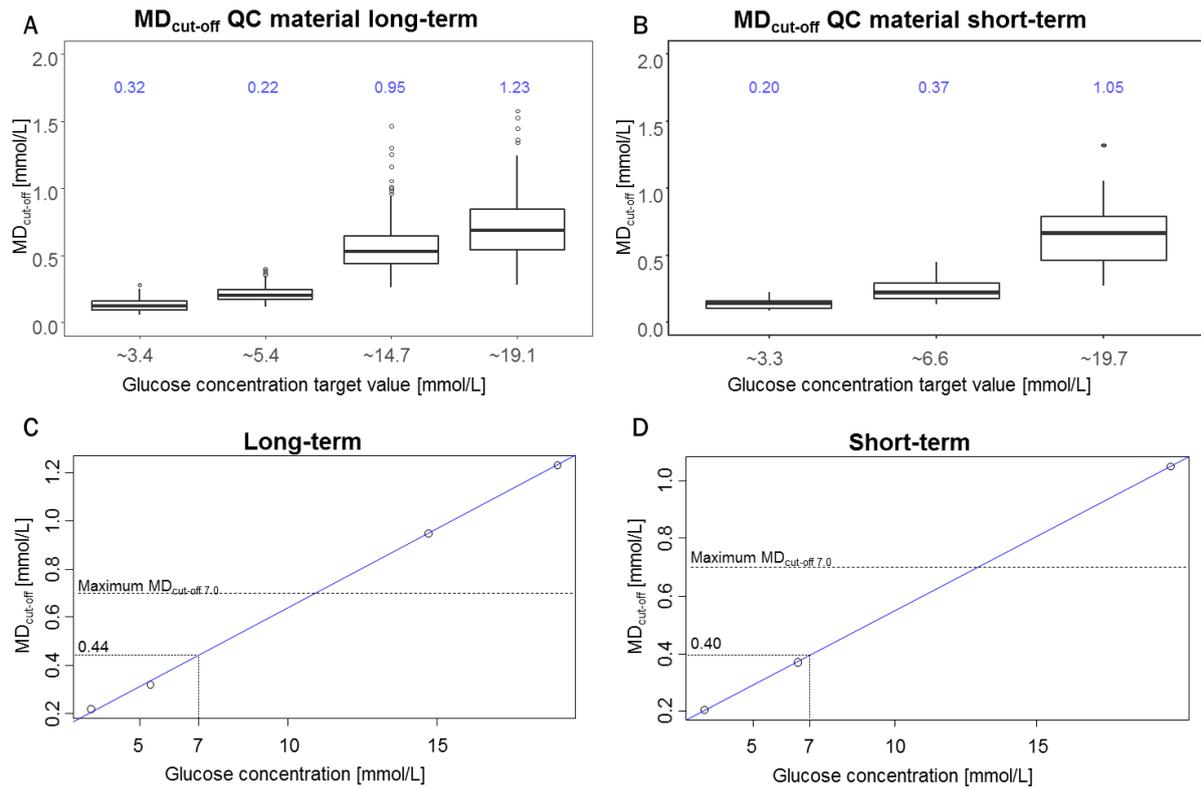


Abbildung 4: Verteilung von $MD_{cut-off}$ bei verschiedenen Konzentrationen.

Verteilung der $MD_{cut-off}$ bei verschiedenen Glukosekonzentrationen (A und B) und lineare Regression (C und D) basierend auf 95% Perzentile $MD_{cut-off}$ (Werte sind über den Boxplots in A und B angegeben). Die lineare Regressionsgerade (blaue durchgezogene Linie) kann verwendet werden, um die MD bei einer ausgewählten Glukosekonzentration, z.B. von 7.0 mmol/L ($MD_{cut-off}$ 7.0), abzuleiten. Das Maximum $MD_{cut-off}$ 7.0 nach DDG ist durch die horizontale gepunktete Linie gekennzeichnet (C und D). Die Daten der Langzeitauswertung von 2009 bis 2018 sind links in A und C dargestellt, die Daten der Kurzzeitdatenauswertung sind rechts in B und D abgebildet.

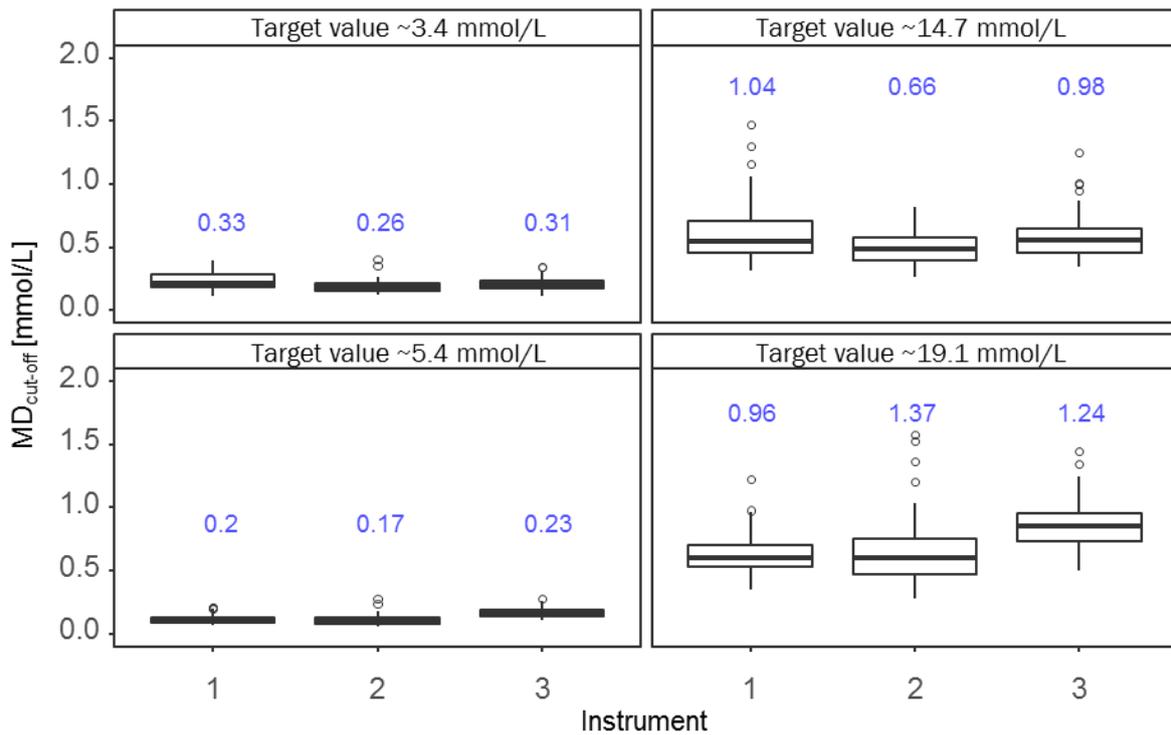


Abbildung 5: Verteilung von $MD_{cut-off}$ in Konzentrationen des verwendeten Kontrollmaterials

Verteilung von $MD_{cut-off}$ in Konzentrationen des verwendeten Kontrollmaterials nach Instrumenten über den gesamten Untersuchungszeitraum von zehn Jahren. Der obere Whisker entspricht dem 1,5-fachen des 3. Quartils, der untere Whisker dem 1,5-fachen des 1. Quartils.

3.2 Kurzzeitauswertung

Die Messungen wurden im Rahmen der Krankenversorgung durchgeführt, zusammen mit Messungen von Patientenproben. Alle Ergebnisse wurden unter Einhaltung der Rili-BÄK gewonnen. Die für die gemäß Rili-BÄK verwendeten QK-Materialien unterschieden sich von den QK-Materialien für die vorliegende Studie. Letztere wurden wie Patientenproben gemessen. Da die regulären QK-Proben nicht Bestandteil des Studiendesigns waren, wurden deren VK separat ermittelt: 2.6% bei einer Glukosekonzentration von 5.2 mmol/L und 1.9% bei 14.7 mmol/L.

Die Evaluation des Bias zwischen den Instrumenten war in dem Kurzzeitexperiment möglich, da das gleiche Material gepoolt wurde und dann über die Instrumente verteilt wurde. Der Bias der Mediankonzentration lag im Bereich von -0.1 bis 0.1 mmol/L, wie in Tabelle 1 gezeigt.

Instrument	Median (mmol/L)	Zentrales 95%Intervall (2,5-97,5 Perzentil) (mmol/L)	Bias (mmol/L)
1	6,5	0,5 (6,2 - 6,7)	-0,1
2	6,7	0,4 (6,5 - 6,9)	0,1
3	6,5	0,5 (6,3 - 6,8)	-0,1
4	6,7	0,5 (6,4 - 7.0)	0,1

Tabelle 1: Bias je Instrument basierend auf der Kurzzeitauswertung.

Arbeitslastprofile

Die Anzahl der stündlichen Glukosemessungen für alle Konzentrationen betrug 1.773, was 88% der theoretisch möglichen ($n = 2.016$) Messungen entspricht. Die fehlenden Messungen waren auf tägliche Wartung und Instandhaltung zurückzuführen.

Das Workload-Profil, einschließlich aller Analyten auf den Vista-Geräten (Abbildung 7A) zeigt einen Hauptpeak bei etwa 9-10 Uhr. Diesem geht ein kleinerer Peak um 6 Uhr mit etwa 4000 - 6000 akkumulierten Tests über den Untersuchungszeitraum von 7 Tagen voraus. Während dieser hohen Arbeitsbelastung wurden zwischen 500 und 800 Labortests pro Stunde durchgeführt. Die geringste Anzahl von Tests wurde in der Zeit von 23 Uhr bis 4 Uhr morgens verzeichnet. Die Arbeitsbelastung wurde ungefähr zu gleichen Teilen auf die Instrumente verteilt, mit Ausnahme von Tag 3, an dem zwei Instrumente den größten Teil der Arbeitsbelastung übernahmen (Abbildung 7B).

Die Unterschiede der Ergebnisse innerhalb und zwischen den Tagen waren sichtbar, aber gering (Abbildung 6) und zeigten keine signifikante Korrelation mit der Arbeitsbelastung.

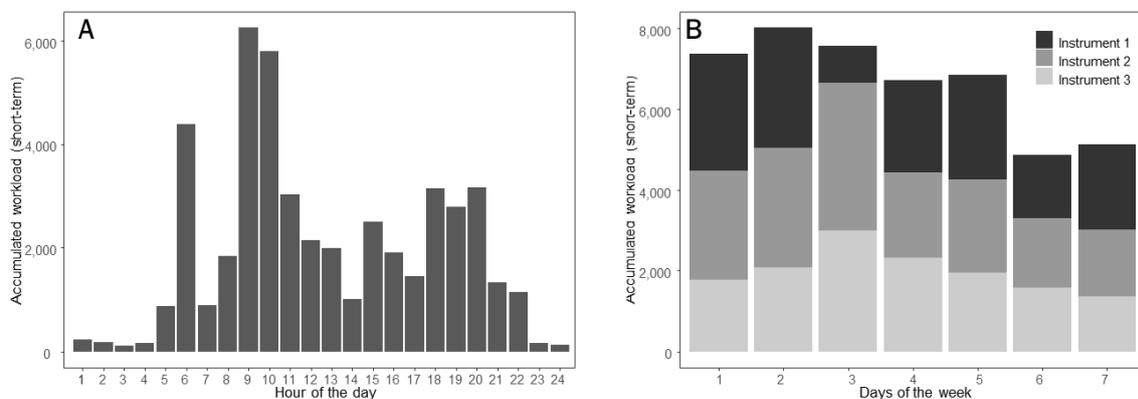


Abbildung 6: Auslastung aller Instrumente in Anzahl der Tests, sortiert nach Tagesszeit und Wochentag

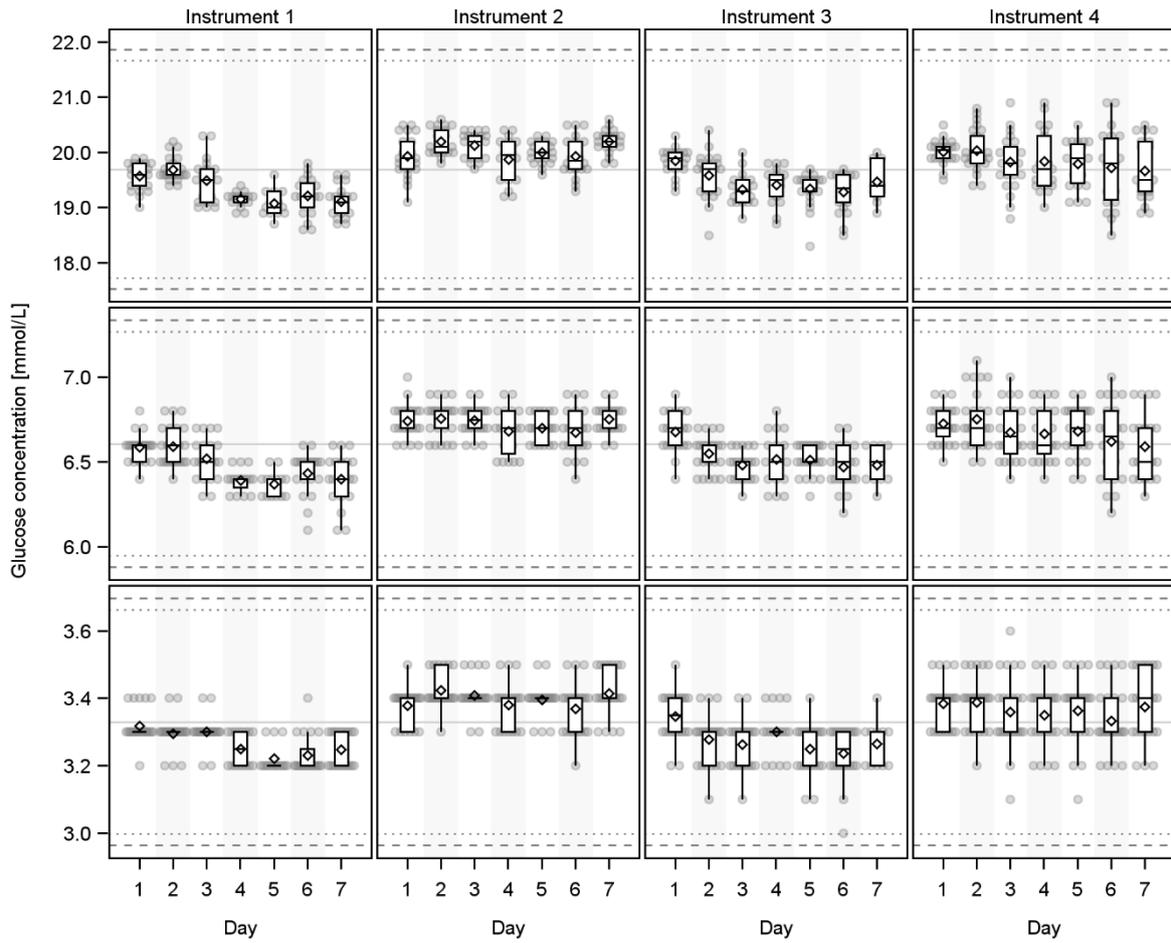


Abbildung 7: Ergebnisse für Glukose

Ergebnisse für Glukose, die vom System über den Untersuchungszeitraum aggregiert angezeigt werden. Horizontale durchgezogene Linien repräsentieren Zielwerte, horizontale gestrichelte Linien stellen Rili-BÄK-Grenzwerte dar (11% für Glukose), und horizontale gepunktete Linien repräsentieren den maximalen MDcut-Off 7.0.

Impräzision

Der $MD_{\text{cut-off}}$ der drei untersuchten Glukosekonzentrationen (3.3 mmol/L, 6.6 mmol/L, und 19.7 mmol/L) im Rahmen der Kurzzeitauswertung ist dargestellt in Abbildung 4B. Die lineare Regression ($y = 0.052x + 0.032$ mmol/L; Abb. 4D) der drei ermittelten $MD_{\text{cut-off}}$ Werte ermöglicht die Berechnung des $MD_{\text{cut-off } 7.0}$, mit 0.40 mmol/L. Der Aufbau des Kurzzeitexperiments erlaubt eine detailliertere Auswertung, beispielsweise die Untersuchung von täglichen Schwankungen innerhalb der Instrumente bei den untersuchten Konzentrationen. Im Kurzzeitexperiment gab es keinen signifikanten Unterschied zwischen dem $MD_{\text{cut-off } 7.0}$ der einzelnen Instrumente bzw. zwischen den Tagen. Der $MD_{\text{cut-off}}$ des Instrumentes 4 (FlexLab) war zwar geringfügig, jedoch signifikant höher als die der anderen Instrumente. Der $MD_{\text{cut-off } 7.0}$ der Kurzzeitauswertung war mit der Langzeitauswertung vergleichbar. Die Messergebnisse werden nach Tag und Gerät detailliert zusammengefasst und ausgewiesen, wie in Abbildung. 6 gezeigt.

4. Diskussion

Für die Beurteilung von Messwerten ist es wichtig, dass sich die behandelnden Ärzte der Messunsicherheit bewusst sind. Bei der Diagnose des Diabetes mellitus hilft dies, um zu entscheiden, ob ein Glukosemesswert verlässlich ober- oder unterhalb einer festgelegten Grenze liegt. Der $MD_{\text{cut-off}}$ beschreibt den kleinsten Abstand zwischen einem Messwert und einem Cut-off, der erforderlich ist, um das Ergebnis als signifikant verschieden vom Cut-off zu beschreiben. Für $k = 2$ (coverage factor) entspricht der $MD_{\text{cut-off}}$ einem Konfidenzintervall von 95% (26). Da die MD in der Einheit des jeweiligen Analyten ausgedrückt wird, muss diese für eine spezifische Konzentration angegeben werden.

Vor dem Hintergrund der Diagnose des Diabetes mellitus konzentriert sich diese Arbeit vor allem auf die Glukosekonzentration von 7.0 mmol/L. Dies ist der diagnostische $MD_{\text{cut-off}}$ -Wert für FPG, wie er von ADA und DDG empfohlen wird. Die entsprechende MD wird als $MD_{\text{cut-off } 7.0}$ angegeben. In dieser Arbeit wurden zwei vergleichende Ansätze zur Ermittlung der Impräzision - ausgedrückt als $MD_{\text{cut-off } 7.0}$ - durchgeführt: Eine Langzeitauswertung der Qualitätskontrollen der internen Qualitätssicherung gemäß Rili-BÄK sowie eine Kurzeitauswertung, bei der über eine Woche stündlich Messungen eines standardisierten Materials durchgeführt wurden.

Der $MD_{\text{cut-off } 7.0}$ wurde mit 0.44 mmol/L in der Langzeitauswertung und mit 0.40 mmol/L in der Kurzeitauswertung bestimmt. Der Zusammenhang zwischen MD und der Glukosekonzentration kann als relative SD, also VK, ausgedrückt werden. Jeder $MD_{\text{cut-off } 7.0}$ auf der linearen Regressionsgeraden in Abbildung 4 repräsentiert denselben VK. Ein $MD_{\text{cut-off } 7.0}$ über oder unter dieser Regressionsgeraden repräsentiert dementsprechend einen höheren bzw. niedrigeren VK. In Laboratorien wird hier i.d.R. der VK verwendet, um die Impräzision von Qualitätskontrollen verschiedener Konzentrationen zu vergleichen.

In beiden Ansätzen blieb der $MD_{\text{cut-off } 7.0}$ unter der empfohlenen Grenze von 0.7 mmol/L. Zudem illustriert und untermauert die Studie das MD-Konzept (20):

1. Es ist zu erwarten, dass 5% aller Ergebnisse außerhalb dieses $MD_{\text{cut-off}}$ liegen. Das zentrale Intervall sollte 95% aller Messergebnisse enthalten. Dieses reicht von 0.4 - 0.5 mmol/L bei einer Konzentration von 6.6 mmol/L in der Kurzzeitauswertung (Tabelle 1). Dieses Ergebnis sollte sich auch im $MD_{\text{cut-off } 7.0}$ aus der Langzeitauswertung widerspiegeln, die dort 0.44 mmol/l betrug.
2. Die Unterschiede zwischen der höchsten und der niedrigsten Glukosekonzentration in der Kurzzeitauswertung bei 6.6 mmol/L variierten zwischen 0.6 und 0.7 mmol/L. Da in dieser Studie $k = 2$ gewählt wurde, ist dies ebenfalls erwartungsgemäß. Würde $k = 3$ gewählt, um einem Konfidenzniveau von 99% zu entsprechen, würde der $MD_{\text{cut-off } 7.0}$ der Langzeitauswertung 0.67 mmol/L betragen, was den Ergebnissen für minimale und maximale Ergebnisse in der Kurzzeitauswertung entspricht.

Darüber hinaus betrug der beobachtete und Rili-BÄK konforme VK bei einer Konzentration von 6.6 mmol/L in der Kurzzeitstudie 2.6%, was einem $MD_{\text{cut-off } 7.0}$ von 0.5 mmol/L entspricht. Dies entspricht wiederum in etwa dem zentralen 95%-Intervall aller Messergebnisse der Kurzzeitauswertung, wie in Tabelle 1 zu sehen.

Die Ergebnisse sind ebenfalls kongruent mit dem $MD_{\text{cut-off}}$, ermittelt aus etwa 21 000 Glukosedoppelmessungen unter Verwendung von Patientenmaterial (27). Dieser betrug 0.38 mmol/L bei 5.0 mmol/L. Dieser Vergleich legt nahe, dass mit Hilfe des verwendeten Kontrollmaterials vergleichbare Resultate erzielt werden wie für den $MD_{\text{cut-off}}$, der mit Patientenmaterial ermittelt wurde, und somit die Übertragung der Ergebnisse auf Glukosekonzentrationsmessungen in Patientenmaterial möglich ist. Angaben hinsichtlich der Impräzision werden vom Hersteller mit $SD = 0.12$ mmol/L bei 4.12 mmol/L und $SD = 0.46$ mmol/L bei 21.02 mmol/L angegeben. Beide Angaben können verwendet werden, um einen $MD_{\text{cut-off } 7.0}$ von 0.36 mmol/L abzuleiten (lineare Regression von $(y = 0,0402x + 0,0742 \text{ mmol/L})$ (28).

Die Impräzisionsangaben des Herstellers beruhen auf einem Protokoll gemäß CLSI/NCCLS EP5-A2, bei dem zweimal täglich zwei Proben je Konzentration gemessen werden. Dies erfolgt über 20 Tage (18). Diese Impräzisionsergebnisse sind niedriger als die erhobenen Ergebnisse, was darauf hinweist, dass das verwendete CLSI/NCCLS EP5-A2 nicht geeignet ist, um einen zuverlässigen $MD_{\text{cut-off } 7.0}$ zu ermitteln.

Der $MD_{\text{cut-off } 7.0}$ wurde aus den Glucosekonzentrationsergebnissen berechnet, die mit drei parallel betriebenen Instrumenten gemessen wurden. Die Unterschiede der $MD_{\text{cut-off } 7.0}$ der einzelnen Instrumente waren in der Langzeitauswertung signifikant, jedoch nicht in der Kurzzeitauswertung. Dies ist ebenfalls ein Hinweis darauf, dass für die Berechnung eines stabilen $MD_{\text{cut-off } 7.0}$ ein angemessen langer Zeitraum besser geeignet ist. Erwartungsgemäß war der $MD_{\text{cut-off } 7.0}$ der einzelnen Instrumente geringfügig niedriger als für das Gesamtsystem, da zusätzlich zur Impräzision auch ein geringfügiger systematischer Unterschied (Bias) zwischen den Instrumenten auftrat und damit die Gesamtvariation, repräsentiert durch die zufällige Probenverteilung durch die Laborstraße (Instrument 4), beeinflusste.

Dennoch erleben klinische Anwender die kombinierte Variabilität aller Instrumente. Die Ergebnisse der Studie zeigen, dass der klinische Anwender selbst bei einer guten Assay Performance jedes einzelnen Instrumentes Unterschiede in den Glukosekonzentrationsergebnissen von bis zu 1.0 mmol/L bei einer Konzentration von 6.6 mmol/L (min. 6.2 mmol/L; max. 7.2 mmol/L) erfahren würde, aufgrund sowohl der Impräzision als auch aufgrund systematischer Abweichungen zwischen den Instrumenten. Daher sollte die $MD_{\text{cut-off } 7.0}$ -Berechnung auf der MD-Verteilung aller Instrumente basieren, die an das automatisierte Verteilungssystem angeschlossen sind.

Die Daten der Langzeitauswertung zeigen eine zuverlässige Stabilität der untersuchten Systeme über ein Jahrzehnt. Dies ist ein Beleg sowohl für die Qualität der Hersteller als auch der medizinischen Laboratorien. Auch auf die Wirksamkeit von Qualitätssicherungssystemen wie die Rili-BÄK mit internen und externen Qualitätskontrollen kann geschlossen werden.

Die Kurzauswertung zeigt zwar geringfügige Variationen der Assay Performance der einzelnen Instrumente, diese lassen sich jedoch nicht mit dem Ausmaß der Auslastung und der Arbeitslast der Instrumente in Verbindung bringen. Obwohl eine hohe Stabilität der Systeme beobachtet wurde, muss angemerkt werden, dass es bei zwei Messungen täglich, also der vorgegebenen QK-Frequenz der Rili-BÄK, möglich ist, hunderte Patientenergebnisse zwischen zwei QKs freizugeben. Eine intensivere QK-Frequenz würde die Anzahl der Patientenergebnisse zwischen zwei QK reduzieren. Doch auch durch eine häufigere Durchführung von QK kann nicht ausgeschlossen werden, dass Ereignisse oder gar Fehlmessungen zwischen zwei QK-Ergebnissen unbemerkt auftreten können. Eine höhere QK-Frequenz verringert jedoch die Anzahl der Proben, die nach einer fehlgeschlagenen QK erneut getestet werden müssen. Es obliegt der Verantwortung des jeweiligen medizinischen Laboratoriums, die Kosten zusätzlicher QK und die Qualität der Analysen für die Patientenversorgung sorgfältig abzuwägen.

Der durch geringfügige Leistungsunterschiede zwischen den Instrumenten verursachten Variabilität kann durch die Ermittlung einer langfristigen $MD_{\text{cut-off } 7.0}$ über alle angeschlossenen Instrumente hinweg Rechnung getragen werden. Die $MD_{\text{cut-off } 7.0}$ war nach etwa 30 unabhängigen QK-Zyklen stabil, was in der Langzeitauswertung nach etwa einem Jahr erreicht wurde, wenn die Ergebnisse aller Instrumente gemeinsam betrachtet wurden.

Der $MD_{\text{cut-off } 7.0}$ der einzelnen Jahre der Langzeitauswertung unterschieden sich, mit Ausnahme des ersten Jahres – dem Einführungsjahr der Instrumente – nicht signifikant. Daher können medizinische Laboratorien QK eines Jahres verwenden, um einen zuverlässigen $MD_{\text{cut-off } 7.0}$ zusammen mit Glukosekonzentrationsergebnissen, die zur Diagnose von Diabetes mellitus verwendet werden, zu berichten. Dies gewinnt besondere Bedeutung, wenn nahe beieinander liegende Cut-offs hohe Anforderungen an die Verlässlichkeit der Ergebnisse stellen (29), um eine richtige Klassifizierung der Patienten zu gewährleisten. Folglich kann durch die kontinuierliche Überwachung der $MD_{\text{cut-off}}$ eine noch zuverlässigere Datenbasis etabliert, Änderungen der Messqualität erfasst sowie der Vergleich verschiedener Messmethoden und Laboratorien unterstützt werden.

4.1 Einschränkungen

Die Studie ist auf einen Instrumententyp (Dimension Vista 1500) und der Teil der Kurzeitauswertung auf nur eine Woche begrenzt. Beeinträchtigungen der Messqualität durch unerwünschte Ereignisse können seltener als wöchentlich auftreten und daher durch das Design dieser Kurzeitauswertung nicht identifiziert werden. Das verwendete Material ist echtem Patientenmaterial sehr ähnlich, jedoch können sogenannte Matrix-Effekte nicht vollständig ausgeschlossen werden. Das Studiendesign ist ausschließlich darauf ausgerichtet, die Impräzision zu erfassen und als MD auszudrücken.

4.2. Ausblick

Weitere Studien mit mehreren Instrumententypen sollten erfolgen. Auch sollte eine weitere Kurzzeitstudie über einen längeren Zeitraum als nur eine Woche angestrebt werden.

In der vorliegenden Arbeit wurde das Konzept zunächst ausschließlich für die Konzentrationsbestimmung Glukose herangezogen. Im Studiendesign wurden weitere Messgrößen erfasst, die nach dem hier vorgestellten Konzept ebenfalls ausgewertet werden können.

6. Zusammenfassung

Die Diagnostik des Diabetes mellitus erfolgt mit Hilfe festgelegter diagnostischer Grenzwerte für bestimmte Biomarker, z.B. für die Plasmaglukosekonzentration. Die DDG empfiehlt in ihrer Praxisempfehlung zur Definition, Klassifizierung und Diagnose von Diabetes mellitus, die Impräzision der Plasmaglukosekonzentrationsbestimmung mit Hilfe der Minimal Difference (MD) anzugeben. Sie leitet sich von der SD ab und beschreibt die kleinste Differenz zwischen zwei Messwerten, die erforderlich ist, um diese als voneinander verschieden zu bezeichnen. Die MD der Glukosekonzentrationsbestimmung soll 0.7 mmol/L bei einer Konzentration von 7.0 mmol/L nicht überschreiten, wenn das Ergebnis für die Diagnose des DM eingesetzt wird. Um verlässliche MD-Werte zu erhalten, wurden in der vorliegenden Studie zwei Ansätze verfolgt:

1. Langzeitauswertung von QK-Daten aus der Krankenversorgung über zehn Jahre Zeitraum Januar 2009 bis Dezember 2018 sowie
2. Kurzzeitauswertung mittels intensiviertem Schema, mit stündlichen Messungen von QK-Material über eine Woche im September 2012.

Die MD der für die Diagnose von DM relevanten Glukosekonzentration wurde in beiden Ansätzen berechnet und verglichen. Die Impräzision der Langzeitauswertung, ausgedrückt als $MD_{\text{cut-off } 7.0}$, betrug 0.44 mmol/L und lag daher deutlich unter dem empfohlenen Grenzwert von 0.7 mmol/L. Stündliche Messungen über einen Zeitraum von einer Woche bestätigen diese Ergebnisse und veranschaulichen das MD-Konzept. Die von QK gemessene Impräzision ist über viele Jahre bemerkenswert stabil. Die derzeitige Impräzisionsbewertung konzentriert sich auf einzelne Instrumente, wohingegen klinisch tätige Ärzte die kombinierte analytische Leistung aller Instrumente, die für eine bestimmten Messgröße in einem Labor verwendet werden, beurteilen. Aus diesem Grund wäre zu empfehlen, den $MD_{\text{cut-off}}$ aus allen Instrumenten und Kontrollzyklen abzuleiten, die für die Verwendung der Patientenversorgung in einem bestimmten medizinischen Labor herangezogen werden. Hier lagen nach ungefähr 30 unabhängigen Kontrollzyklen, bzw. einem Jahr, stabile Ergebnisse für $MD_{\text{cut-off } 7.0}$ vor. Die Einrichtung einer kontinuierlichen Überwachung der $MD_{\text{cut-off}}$ kann die traditionelle Qualitätssicherung ergänzen.

7. Literaturverzeichnis

1. Petersmann A, Müller-Wieland D, Müller UA, Landgraf R, Nauck M, Freckmann G, et al. Definition, Klassifikation und Diagnostik des Diabetes mellitus. *Diabetol Stoffwechs* 2019;14:S111-8.
2. Hien P, Claudi-Böhm S, Krämer C, Kohlhas K. *Diabetes-Handbuch*. 7 ed. Heidelberg: Springer-Verlag; 2013.
3. Kerner W, Brückel J. Definition, Klassifikation und Diagnostik des Diabetes mellitus. *Diabetol Stoffwechs* 2014;9: S96-9.
4. WHO | Definition and diagnosis of diabetes mellitus and intermediate hyperglycaemia. WHO. 2006.
5. Shaw JE, Sicree RA, Zimmet PZ. Global estimates of the prevalence of diabetes for 2010 and 2030. *Diabetes Res. Clin. Pract.* 2010; 87:4–14.
6. Zimmet P, Alberti KG, Shaw J. Global and societal implications of the diabetes epidemic. *Nature* 2001; 414:782-7.
7. 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;15:539-53.
8. Zimmet P, Alberti KG, Shaw J. Global and societal implications of diabetes epidemic. *Nature*. 2001; 414:782-7.
9. Wild SH, Roglic G, Green A, Sicree R, King H. Global prevalence of diabetes: estimates for the year 2000 and projections for 2030: response to Rathman and Giani. *Diabetes Care* 2004; 27: 2569-70.
10. American Diabetes Association. *Diagnosis and Classification of Diabetes Mellitus*. *Diabetes Care* 2012; 36: S67–74.
11. Expert Committee on the Diagnosis and Classification of Diabetes Mellitus. Report of the expert committee on the diagnosis and classification of diabetes mellitus. *Diabetes Care*. 2003; 26: S5–20.
12. Rett K. The relation between insulin resistance and cardiovascular complications of the insulin resistance syndrome. *Diabetes Obes Metab.* 1999;1: S8-16.
13. King H, Aubert RE, Herman WH. Global burden of diabetes 1995-2025. *Diabetes Care* 1998; 21: 1414-31.
14. Forouhi NG, Wareham NJ (). *Epidemiology of diabetes*. *Medicine* (Abingdon, England: UK ed). 2014; 42: 698-702.
15. World Health Organization. Use of glycated haemoglobin (HbA1c) in the diagnosis of diabetes mellitus: abbreviated report of a WHO consultation. WHO guidelines approved by the Guidelines Review Committee. Geneva: World Health organization; 2011.
16. Marathe PH, Gao HX, Close KL. American Diabetes Association Standards of Medical Care in Diabetes 2017. *J Diabetes* 2017;9:320-4.
17. Müller-Wieland D, Petersmann A, Nauck M, Heinemann L, Kerner W, Müller UA, et al. Definition, classification and diagnosis of diabetes mellitus. *Diabetol Stoffwechs*. 2016; 11: S78-81.

18. Joint Committee for Guides in Metrology (JCGM). Evaluation of measurement data-guide to the expression of uncertainty in measurement. JCGM 100:2008. <https://www.bipm.org/en/publications/guides/gum.html>.
19. Büttner J, Broth R, Broughton PM, Bowyer RC. International Federation of Clinical Chemistry. Committee on standards. Expert panel on nomenclature and principles of quality control in clinical chemistry. IFCC recommendation. Quality control in clinical chemistry control. J Clin Chem Clin Biochem 1980; 18 :535-41.
20. Clinical and Laboratory Standards Institute (CLSI). Expression of measurement uncertainty in laboratory medicine; approved guideline. 2012: Document EP29.
21. Parvin CA. New insight into the comparative power of quality-control rules that use control observations within a single analytical run. Clin Chem. 1993; 39: 440-7.
22. Westgard JO, Barry PL, Hunt MR, Groth T. A multi-rule Shewhart chart for quality control in clinical chemistry. Clin Chem. 1981;27: 493-501.
23. Neufassung der Richtlinie der Bundesärztekammer zur Qualitätssicherung laboratoriumsmedizinischer Untersuchungen Rili-BÄK Dtsch. Arztebl. 2019;116:A-2422 / B-1990 / C-1930
24. Kunst A, Drager B, Ziegenhorn J. UV methods with hexokinase and glucose-6-phosphate dehydrogenase. In: Bergemeier H, edizor. Methods of enzymatic analysis. Deerfield: Verlag Chemie; 1983:163-72.
25. Badrick T, Cervinski M, Loh TP. A primer on patient-based quality control techniques. Clin Biochem. 2019;64:1-5.
26. Petersen PH, Ricos C, Stockl D, Libeer JC, Baadenhuijsen H, Fraser C, et al. Proposed guidelines for the internal quality control of analytical results in the medical laboratory. Eur J Clin Chem Clin Biochem. 1996; 34: 983-99.
27. Petersmann A, Wasner C, Nauck M, Kallner A. Frequency of extreme difference and clinical performance of glucose concentration measurements judged from 21,000 duplicate measurements. Clin Chem. 2013;59:998-1000.
28. Siemens Healthcare Diagnostics. Dimensions Vista System reagent cartridge, glucose. Ref. K1039. Newark: Siemens Healthcare Diagnostics; 2017.
29. D´Orazio P, Burnett RW, Fogh-Andersen N, Jacobs E, Kuwa K, Kulpmann WR, et al. Approved IFCC recommendation on reporting results for blood glucose (abbreviated). Clin Chem. 2005;51:1573-6.

8. Abbildungsverzeichnis

Abbildung 1: Beziehung zwischen VK und Minimal Difference (MD) an den diagnostischen Cut-offs der Glukosekonzentration, die zur Diagnose von DM empfohlen werden.....	6
Abbildung 2: Dimension Vista angebunden an die Laborstraße FlexLab (Siemens Healthcare GmbH, Eschborn, Deutschland) im Institut für Klinische Chemie und Laboratoriumsmedizin der Universitätsmedizin Greifswald.....	9
Abbildung 3: Hexokinase-Glukose-6-Phosphat-Dehydrogenase-Methode.....	10
Abbildung 4: Verteilung von $MD_{\text{cut-off}}$ bei verschiedenen Konzentrationen.....	16
Abbildung 5: Verteilung von $MD_{\text{cut-off}}$ in Konzentrationen des verwendeten Kontrollmaterials.....	17
Abbildung 6: Auslastung aller Instrumente in Anzahl der Tests, sortiert nach Tagesszeit und Wochentag.....	19
Abbildung 7: Ergebnisse für Glukose.....	20

9. Tabellenverzeichnis

Tabelle 1: Bias je Instrument basierend auf der Kurzeitauswertung.....	18
--	----

10. Formelverzeichnis

Formel 1: Gleichung (1)-(5)11

11. Wissenschaftliche Publikation

Sandra Keutmann, Stephanie Zylla, Mathilde Dahl, Nele Friedrich, Rüdiger Landgraf, Lutz Heinemann, Anders Kallner, Matthias Nauck, Astrid Petersmann

Measurement Uncertainty Impacts Diagnosis of Diabetes Mellitus: Reliable Minimal Difference of Plasma Glucose Results

Diabetes Therapy (2020) 11:293-303

<https://doi.org/10.1007/s13300-019-00740-w>



Measurement Uncertainty Impacts Diagnosis of Diabetes Mellitus: Reliable Minimal Difference of Plasma Glucose Results

Sandra Keutmann · Stephanie Zylla · Mathilde Dahl · Nele Friedrich ·
Rüdiger Landgraf · Lutz Heinemann · Anders Kallner ·
Matthias Nauck · Astrid Petersmann

Received: August 22, 2019
© The Author(s) 2019

ABSTRACT

Introduction: The diagnosis of diabetes mellitus is based on suitable cut-off values of specific biomarkers, such as the concentration of glucose in plasma. The German Diabetes Association has very recently published a clinical practice guideline on the definition, classification and diagnosis of diabetes mellitus that recommends measurements of plasma glucose concentration have an imprecision defined as a minimal difference (MD) of at a fasting plasma

glucose concentration of 7.0 mmol/L. To obtain reliable values for the MD, we investigated long-term and short-term measurement uncertainty.

Methods: The imprecision was determined by two approaches: (1) a long-term dataset with imprecision based on the Guideline of the German Medical Association on Quality Assurance in Medical Laboratory Examinations (Rili-BAEK), in a medical laboratory operating 24/7, using internal quality control (IQC) data for four concentrations during a 10-year period; and (2) a detailed short-term dataset with imprecision assessed by hourly measurements of control materials. These datasets were used to calculate the MD cut-off ($MD_{\text{cut-off}}$) as: $MD_{\text{cut-off}} = k \times SD = 2 \times SD$, where SD is the standard deviation and $k = 2$ represents a confidence level of 95%.

Results: The $MD_{\text{cut-off}}$ of ≤ 0.7 mmol/L at a fasting plasma glucose concentration of 7.0 mmol/L ($MD_{\text{cut-off } 7.0}$) for the long-term and

Enhanced Digital Features To view enhanced digital features for this article go to <https://doi.org/10.6084/m9.figshare.10790468>.

Electronic supplementary material The online version of this article (<https://doi.org/10.1007/s13300-019-00740-w>) contains supplementary material, which is available to authorized users.

S. Keutmann · S. Zylla · M. Dahl · N. Friedrich ·
M. Nauck · A. Petersmann (✉)
Institute of Clinical Chemistry and Laboratory
Medicine, University Medicine Greifswald,
Greifswald, Germany
e-mail: astrid.petersmann@med.uni-greifswald.de

S. Zylla · N. Friedrich · M. Nauck
DZHK (German Centre for Cardiovascular
Research), Partner Site Greifswald, University
Medicine Greifswald, Greifswald, Germany

R. Landgraf
German Diabetes Foundation (DDS), Munich,
Germany

L. Heinemann
Science-Consulting in Diabetes GmbH, Neuss,
Germany

A. Kallner
Department of Clinical Chemistry, Karolinska
University Hospital, Stockholm, Sweden

Present Address:
A. Petersmann
Institute of Clinical Chemistry, University Medicine
Göttingen, Göttingen, Germany

the short-term approaches were 0.44 and 0.40 mmol/L, respectively. The $MD_{\text{cut-off } 7.0}$ from both approaches was therefore below the recommended value of 0.7 mmol/L. It was noted that the variability in performance within and between instruments can be covered by reporting the long-term $MD_{\text{cut-off } 7.0}$ across all connected instruments. In this study, stable results for the $MD_{\text{cut-off } 7.0}$ were obtained after 1 year.

Conclusion: Imprecision as measured by IQC data is remarkably stable over many years of operation. Current imprecision assessment usually focuses on only single instruments, whereas clinicians perceive the measurement as the result of the combined analytical performance of all instruments used for a certain assay. In the clinical setting, the MD may be a more useful measure of imprecision, and we suggest deriving the $MD_{\text{cut-off}}$ combined from all instruments and control cycles that are used in the patient care setting for a given analyte.

Keywords: Assay performance; Coefficient of variation; Imprecision; Internal quality control; Minimal difference; Rili-BAEK; Westgard rules

Key Summary Points

All laboratory measurements are subject to measurement uncertainty, which can be described using the coefficient of variation (%CV). Despite the pervasive presence of high measurement uncertainty and the enormous impact it may have on the diagnosis of diabetes mellitus (DM), measurement uncertainty has not received appropriate attention by clinical users.

Our aim was to create awareness of the fact that the magnitude of measurement uncertainty especially impacts the diagnosis of DM when cut-off values are used. We calculated the recommended “minimal difference (MD)”, a value recently developed for this purpose, based on detailed short-term and long-term (10 years) internal quality control data for state-of-the-art glucose concentration measurements.

The MD represents the smallest difference between two values that is statistically significant, taking the specific measurement uncertainty into account; it is based on the standard deviation and expressed in the unit of the measurand to enable easy use by clinicians.

At the fasting plasma glucose cut-off of 7.0 mmol/L (126 mg/dL), which represents a frequently used cut-off value for the diagnosis of DM, the MD cut-off ($MD_{\text{cut-off } 7.0}$) for the investigated assay was 0.44 mmol/L and remained below the recommended maximum MD of 0.7 mmol/L.

State-of-the-art glucose concentration measurements with low imprecision are associated with a considerable MD and therefore may impact cut-off points for diagnosing diabetes. It is important to note that imprecisions allowed for by guidelines or legal requirements are considerably higher than those observed in this study. Consequently, imprecision in glucose assays should also be reported as MD to enable clinical users to consider imprecision when relying on the results of a glucose assay to diagnose DM.

INTRODUCTION

The diagnosis of diabetes mellitus (DM) is based on suitable cut-off values for specific biomarkers, such as the plasma glucose concentration. The World Health Organization (WHO), American Diabetes Association (ADA) and German Diabetes Association (DDG) recommend assaying samples of venous plasma, with a cut-off for fasting plasma glucose (FPG) of ≥ 7.0 mmol/L (≥ 126 mg/dL) and for 2-h plasma glucose of ≥ 11.1 mmol/L (200 mg/dL) after ingestion of 75 g of glucose in the oral glucose tolerance test [1–3]. According to WHO, impaired fasting glucose (IFG) is diagnosed when the plasma glucose

level is between 6.1 and 6.9 mmol/L (110–125 mg/dL); according to ADA and DDG, the interval in plasma glucose level for IFG diagnosis is 5.6–6.9 mmol/L (100–125 mg/dL) [3].

A Laboratory Result is Complete only when Reported Together with the Associated Measurement Uncertainty

Laboratory reports are often treated like a bank account statement: the value that is given on the statement is understood to be the actual amount of money in the bank account, and deviation by a few cents more or less is not acceptable. In contrast to the values on bank account statements, however, laboratory measurements are always subject to measurement uncertainty, and a few “cents” more or less in the results may represent the same value, depending on the magnitude of the associated uncertainty. Although it is generally accepted that a measurement result is only complete if the associated uncertainty is reported together with the measurement result, as stated in “The Guide to the Expression of Uncertainty in Measurements” [4], the uncertainty is usually not reported to the clinical end-user. Medical laboratories commonly report imprecision as the coefficient of variation (%CV) [5]. The %CV, as given in Eq. (1), may be difficult to immediately translate to a difference between two consecutive results or between a result and a cut-off value. It is important that the end-user—the clinician—takes the measurement uncertainty into account when comparing a laboratory result to a recommended cut-off value, such as, for example, when diagnosing or treating DM. Therefore, the uncertainty expressed in the units of the measurand, i.e. the minimal difference (MD), would be a more useful metric.

Description of Minimal Difference

The DDG recently published a clinical practice guideline on the definition, classification and diagnosis of DM that recommends minimum requirements for the imprecision of glucose

concentration measurements, expressed as the MD [2]. The MD is not a novel metric and has been described previously—for example, in the approved guideline of the Clinical and Laboratory Standards Institute (CLSI) EP29-A on “Expression of Measurement Uncertainty in Laboratory Medicine” [6]. The MD represents the smallest difference between two values that is statistically significant, taking the specific measurement uncertainty into account. It is based on the standard deviation (SD) and is expressed in the unit of the measurand so that clinicians are able to apply it directly to the result [7].

At the FPG cut-off of 7.0 mmol/L, the DDG recommends a MD cut-off ($MD_{\text{cut-off}}$) of ≤ 0.7 mmol/L ($MD_{\text{cut-off } 7.0}$) for FPG measurements used for the diagnosis of DM. At this glucose concentration of 7.0 mmol/L, a SD = 0.35 mmol/L and a MD of 0.7 mmol/L (Eq. 5) correspond to a %CV of $\leq 5\%$. A %CV of 6% would correspond to 0.84 mmol/L, i.e. any value between 6.16 and 7.84. Applied to a concentration of 5.6 mmol/L the interval would be 4.93–6.27 mmol/L. An analytical %CV of this magnitude, or larger, would therefore not differentiate between diabetes and non-diabetes. This is illustrated by the intersection and overlap of the “funnels” in Fig. 1.

Current quality assurance regulations, such as the commonly practiced Westgard control rules [8] or the compulsory Guideline of the German Medical Association on Quality Assurance in Medical Laboratory Examinations (RiliBAEK) [9], usually require at least two internal quality control (IQC) samples of different concentrations per assay, applied at least once each per instrument per day. If the daily requirements of the control system are met, no further action need be taken and the results are released. It is thus assumed that results obtained between two accepted control samples meet the quality criteria [5].

Here we report our assessment of the analytical imprecision of glucose concentration measurements on a continuously operated state-of-the-art measurement system. Glucose measurements were performed on three individual instruments connected to an automatic random sample distribution system. Two

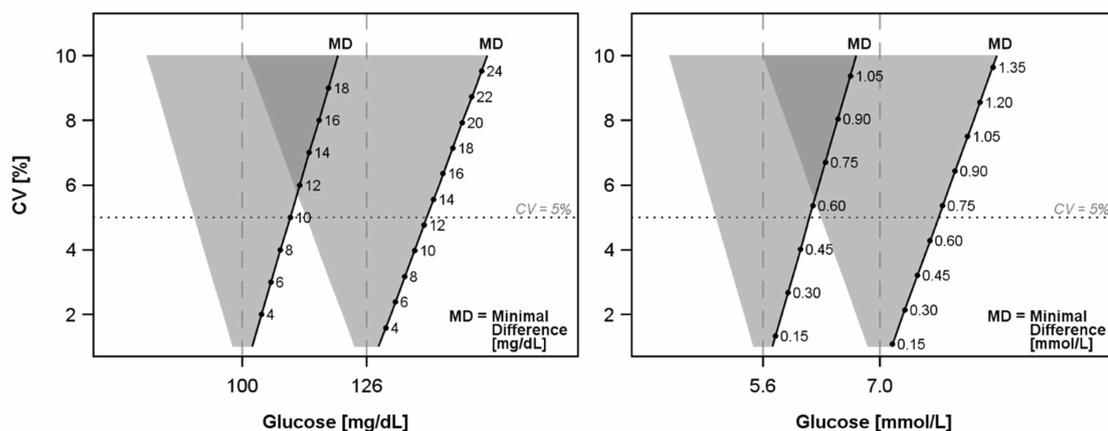


Fig. 1 Relation between coefficient of variation (%CV) and minimal difference (MD) at glucose concentration cut-offs used for the diagnosis of diabetes mellitus. The width of the funnels at each %CV is the MD and can be read on the oblique scale. The vertical dotted lines are the cut-off values. The overlapping areas above the intersection

of the funnels correspond to %CVs where the difference between 5.6 and 7.0 mmol/L glucose cannot be ascertained. Figure is reproduced with permission from M Nauck, A Petersmann, D Müller-Wieland, et al. Definition, Klassifikation und Diagnostik des Diabetes mellitus. Diabetol Stoffwech. 2018;13

approaches were chosen: (1) long-term imprecision calculated from IQC data obtained in a regular patient care setting between 2009 and 2018; and (2) detailed short-term imprecision assessed hourly through the measurement of additional control materials during a 1-week period. The $MD_{\text{cut-off}}$ of both approaches was calculated and compared for the FPG cut-off value of 7.0 mmol/L (126 mg/dL) for the diagnosis of DM.

The aim of the study was to provide reliable imprecision data for glucose concentration measurements and translate this imprecision into a clinical useable term, namely the MD. This information is important in view of the fact that each measurement result is subject to imprecision, and clinicians should be aware of the magnitude of this imprecision, especially when using cut-offs for diagnostic purposes, such as in the diagnosis of DM.

METHODS

All measurements were conducted at the Institute of Clinical Chemistry and Laboratory Medicine of the University Medicine of

Greifswald, which operates 24 h, 7 days a week. The long-term study was performed between January 2009 and December 2018. The measuring systems were used for patient care and complied with the Rili-BAEK guideline. The plasma glucose concentrations were measured using calibrators, reagents and instruments (hexokinase-glucose-6-phosphate dehydrogenase method [10]) on three Dimension Vista 1500 (instruments 1–3) that were connected by a laboratory automation system (all from Siemens Healthcare GmbH, Eschborn, Germany). The automation system was changed from StreamLab to FlexLab (instrument 4; Siemens Healthcare GmbH) in 2012. Both automation systems distributed the samples to analytical Dimension Vista instruments 1–3. The IQC material Tru Liquid Monitrol (Thermo Fisher Scientific, Schwerte, Germany) was used from 2009 until December 2013, after which time it was changed to Liquid Assayed Multiquel (Bio-Rad Laboratories, Munich, Germany).

No patient material was used for this study; therefore, the study did not need to be registered, and it did not require the approval of an institutional ethics committee.

Minimal Difference

The MD is based on the SD and is expressed in the unit of the measurand to enable clinicians to apply it directly to the result [7]. If a coverage factor (k) of 2 is used, the MD corresponds to a confidence level of about 95%. If two consecutive measurements are compared, the measurement uncertainty of both must be considered (Eq. 2). If measurement uncertainty is the same for each measurement, the equation can be simplified, as given in Eq. (3). If a measurement result is compared to a fixed cut-off value that is normally considered to be devoid of an uncertainty, the MD can be simplified to the $MD_{\text{cut-off}}$ as calculated by Eq. (4), which in turn can be simplified to Eq. (5).

$$\% CV = \frac{SD}{\bar{X}} \times 100; SD = \frac{\bar{X} \times \%CV}{100} \quad (1)$$

$$MD = k \times \sqrt{SD_1^2 + SD_2^2} \quad (2)$$

If $SD_1 = SD_2$:

$$MD = k \times \sqrt{2 \times SD^2} = 2 \times SD \times \sqrt{2} \quad (3)$$

If $SD_2 = 0$:

$$MD_{\text{cut-off}} = k \times \sqrt{SD_1^2}, \quad (4)$$

which can be simplified to:

$$MD_{\text{cut-off}} = k \times SD = 2 \quad (5)$$

In Eq. (1) to (5), \bar{X} is the mean and $k = 2$, which represents a confidence level of 95%.

Long-Term Experiment

Imprecision

Documented IQC data from the laboratory information system (LIS) were used retrospectively for the study period of 10 years. Control cycles were evaluated on a monthly basis, separately for each instrument, IQC level and much of the control material. Cycles with fewer than 15 IQC results were excluded.

Short-Term Experiment

Workload

There was a total of 2016 possible measurement time points for all four instruments. The actual number of measurements was documented in order to correctly handle interruptions for service and maintenance. The overall workload, i.e. the total number of measurements per hour of the day, was retrieved retrospectively from the LIS for each of the instruments.

Imprecision

Control material was measured hourly over a period of 1 week in September 2012. Samples were directly introduced into each Dimension Vista instrument (1–3) or randomly through FlexLab, imitating the general work flow of patient samples. The latter procedure represents instrument 4. Performance of the instruments and the measurements were independently monitored with separate IQC materials according to the rules in the Rili-BAEK. Acceptance of these rules was the inclusion criterion for the measurements of both the patient care and study samples. Short-term imprecision results were performed as if they were patient samples and collected from the LIS retrospectively; they were not reported separately, nor used in the Rili-BAEK-based IQC system. Short-term control cycles were evaluated on a daily basis from the hourly measurements, separately for each instrument, IQC level and lot number.

Three levels of control material with concentrations of about 3, 7 and 20 mmol/L (Bio-Rad Liquid Assayed Multiqual; lot no. 45,631-1 through -3; Bio-Rad Laboratories) were used. For each level, a sufficient volume of each material to cover the needs for 1 day was thawed and pooled for each concentration and then aliquoted into barcode-labeled and capped tubes suitable for use in the instruments. The aliquoted material was stored at 2–5 °C until use (maximum storage time 24 h).

Calculations and Statistics

Calculation of the %CV and MD, including descriptive statistics and boxplots, was

performed using R (version 3.5.0; release 23 March 2018; <https://www.r-project.org/>). Statistical differences between instruments and years were calculated using analysis of variance, and Tukey's HSD test was used as the post hoc test. Frequency graphs were designed using SAS version 9.4 software (SAS Institute Inc., Cary, NC, USA). The MD was calculated using Eq. (5), i.e. under the assumption that the cut-off was not liable to an uncertainty. All calculations are given in Electronic Supplementary Material (ESM) S1.

In the long-term approach, the MDs were calculated monthly for each instrument using the IQC results and then cumulated as a moving average to obtain a reliable value [11]. In the short-term approach, the MDs were calculated daily for each instrument. Equations (1) and (5) were applied for both approaches. All MDs were summarized in boxplots for the whole study period and also reported in detail by years and instruments, respectively.

The 95th percentiles of the $MD_{\text{cut-off}}$ distributions were used to obtain reliable values for the MD suitable for the diagnosis of DM. The $MD_{\text{cut-off}}$ from the different concentrations tested in this study were used to calculate a linear equation that would allow the $MD_{\text{cut-off}}$ for other concentrations, such as the $MD_{\text{cut-off } 7.0}$, to be estimated.

Since the material used in the short-term experiment was pooled and then distributed to the instruments, an evaluation of the bias among the instruments was possible. This bias was calculated separately for each instrument by subtracting the mean $MD_{\text{cut-off}}$ of each instrument from the median $MD_{\text{cut-off}}$ of all instruments.

Compliance with Ethics Guidelines This article does not contain any studies with human participants or animals performed by any of the authors.

RESULTS

Long-Term Experiment

Imprecision

Long-term glucose measurement imprecision was calculated from the IQC samples tested

during patient care under routine conditions. The calculated values converged and reached a stable level after 30 observations. All results were obtained following the Rili-BAEK and are summarized in Fig. 2. The $MD_{\text{cut-off } 7.0}$ was calculated from the linear equation $y = 0.065x - 0.013$, which was derived from four concentrations (3.2, 5.5, 15.9, and 19.2 mmol/L) and determined to be 0.44 mmol/L (Fig. 2c). The $MD_{\text{cut-off } 7.0}$ differed marginally but significantly between the instruments (ESM S2, S3). This tendency was also found for a few IQCs, which were found to be higher in 2008 than in any other year; after a change in IQCs in 2014, the performance was stable throughout all years, i.e. there was no significant difference in the $MD_{\text{cut-off } 7.0}$ (ESM S2).

Short-Term Experiment

All measurements were performed together with measurements taken for patient care under routine conditions. All results were compiled using Rili-BAEK. Since the regular IQC samples were not part of the study design in the short-term approach, the %CVs from the routine IQC were taken separately: 2.6 and 1.9% at a glucose concentration of 5.2 and 14.7 mmol/L, respectively.

Evaluation of the bias among the instruments was possible since the sample material used in the short-term experiment was pooled and then distributed across the instruments. The bias of the median concentration ranged from -0.1 to 0.1 mmol/L. Relative contributions of each instrument to the workload are given in Fig. 3.

Workload Profiles

A total of 1773 hourly glucose measurements for all concentrations was available for the analysis, which corresponds to 88% of all possible ($n = 2016$) measurements. The measurements that were missing were due to daily service and maintenance procedures. Examination of the daily workload profile, including all analytes provided for regular patient care on the Vista instruments (Fig. 3a), revealed one main peak at about 9–10 a.m., preceded by a smaller

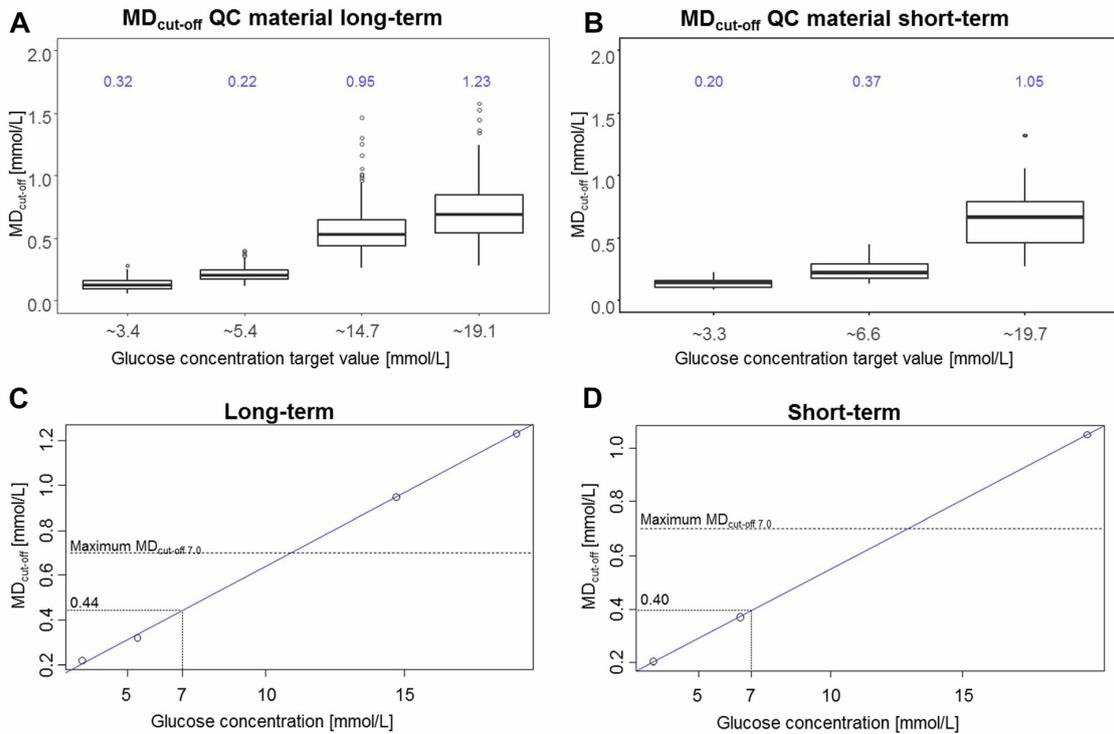


Fig. 2 Distribution of the MD cut-offs ($MD_{cut-off}$) at different concentrations of glucose (**a**, **b**) and linear regressions (**c**, **d**) based on the 95th percentile $MD_{cut-off}$ (values are given above the boxplots in **a** and **b**). The linear regression (blue solid line) can be used to derive the MD at a plasma glucose cut-off value of 7.0 mmol/L ($MD_{cut-off 7.0}$)

as indicated by the solid black line. The maximum $MD_{cut-off 7.0}$ according to the German Diabetes Association (DDG) is marked by the horizontal dotted line. **a**, **c** Long-term internal quality control (IQC) data from 2009 to 2018, **b**, **d** short-term data from hourly measurements of QC material during 1 week

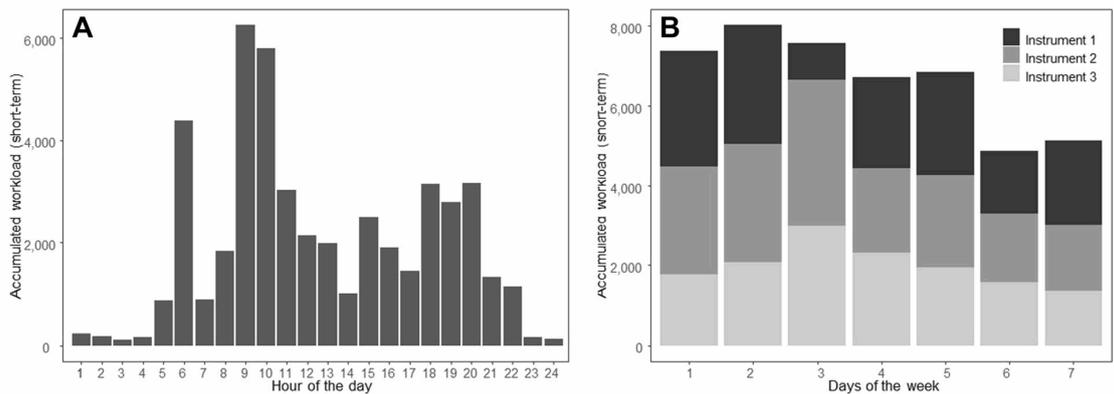


Fig. 3 Workload of all instruments in number of tests according to hour of the day (**a**) and day of the week (**b**)

peak at 6 a.m., with a range of approximately 4000–6000 accumulated tests over the study period of 7 days. Between 500 and 800 laboratory tests were performed during these heavy workload hours. The lowest number of tests was recorded between 11 p.m and 4 a.m. The workload was shared approximately equally between the instruments, except for day 3 when two instruments took over most of the workload (Fig. 3b). Variations in results between and within days were noticeable but small (ESM S4) and showed no dependence on the workload.

Imprecision

The $MD_{\text{cut-off}}$ of the three investigated glucose concentrations (3.3, 6.6 and 19.7 mmol/L) are summarized in Fig. 2. The linear regression ($y = 0.052x + 0.032$; Fig. 2d) of the three obtained $MD_{\text{cut-off}}$ values enabled the calculation of the $MD_{\text{cut-off } 7.0}$, which was 0.40 mmol/L.

The design of the short-term experiment permitted a more detailed evaluation, such as reporting daily variations within instruments for the investigated concentrations. In the short-term approach there was no significant difference between the $MD_{\text{cut-off } 7.0}$ of the individual instruments or between days. The $MD_{\text{cut-off } 7.0}$ of instrument 4 (Flexlab) was slightly but significantly higher than that of the other instruments. The $MD_{\text{cut-off } 7.0}$ of the short-term experiment was comparable to that from the long-term experiment. The measurements results are summarized and reported in detail by day and instrument in ESM S4.

DISCUSSION

Clinicians should be aware of measurement uncertainty, such as, for example, when comparing the result of a glucose concentration measurement to a cut-off value used for the diagnosis of DM.

The $MD_{\text{cut-off}}$ is a metric denoting the smallest analytical difference between a measurement and a cut-off that can be regarded as statistically significant at a 95% confidence level—provided a coverage factor (k) of 2 is applied [12]. Since the $MD_{\text{cut-off}}$ is given in the

unit of the analyte it can only be reported for a specific concentration. As the subject of our study was the $MD_{\text{cut-off}}$ for the diagnosis of DM we focused on glucose concentrations at 7.0 mmol/L, which is the diagnostic cut-off value for FPG as recommended by ADA and DDG and assessed long-term and short-term imprecision expressed as $MD_{\text{cut-off } 7.0}$ for glucose concentration measurements.

The $MD_{\text{cut-off } 7.0}$ for glucose in the long-term and the short-term approaches were 0.44 and 0.40 mmol/L, respectively. The relation between the $MD_{\text{cut-off}}$ and the glucose concentration represents the relative SD, i.e. the %CV. Any $MD_{\text{cut-off}}$ on the linear regression line in Fig. 2 represents the identical %CV. A $MD_{\text{cut-off}}$ above and below this regression line represents a higher and a lower %CV, respectively. In the laboratory setting, %CV can also be used to compare imprecisions of different quality control concentrations.

In our study, the $MD_{\text{cut-off } 7.0}$ remained below the recommended limit of 0.7 mmol/L in both approaches. The results of the study illustrate the MD concept [6] as follows:

1. It can be expected that 5% of all glucose concentration measurements fall outside this $MD_{\text{cut-off}}$. The central 95% interval of all measurement results in the short-term study was found to be 0.4–0.5 mmol/L at a glucose concentration of 6.6 mmol/L (ESM S4). This interval should resemble the $MD_{\text{cut-off } 7.0}$ derived from the long-term experiment, which in turn was shown to be 0.44 mmol/L.
2. The differences between the highest and lowest glucose concentrations of study samples in the short-term experiment at 6.6 mmol/L varied between 0.6 and 0.7 mmol/L. Since a coverage factor $k = 2$ was chosen in this study, this result is reasonable. If the coverage factor (k) is increased to 3 to correspond to a 99% level of confidence, the $MD_{\text{cut-off}}$ from the long-term study would be 0.67 mmol/L, which also matches our findings for minimum and maximum results in the short-term study.

In addition, routine Rili-BAEK IQC in the short-term approach, which relied on separate

control material, confirmed that the %CV at a concentration of 6.6 mmol/L was 2.6%, which corresponds to a $MD_{\text{cut-off}}$ of 0.5 mmol/L. This also resembles the central 95% interval of all measurement results in the short-term study.

The results of the present study are also in line with the $MD_{\text{cut-off}}$ of 0.38 mmol/L at 5.0 mmol/L calculated from approximately 21,000 duplicate measurements of glucose in plasma collected from patients instead of from IQC material [13]. This result demonstrates that the control material used gives $MD_{\text{cut-off}}$ results comparable to those determined for plasma samples from patients and therefore allows the $MD_{\text{cut-off}}$ to be extrapolated to glucose concentration measurements of plasma samples from patients. Imprecision performance data claimed by the manufacturer are a $SD = 0.12$ mmol/L at 4.12 mmol/L and a $SD = 0.46$ mmol/L at 21.02 mmol/L, both of which can be used to derive a $MD_{\text{cut-off } 7.0}$ of 0.36 (linear regression of $y = 0.0402x + 0.0742$) [14]. The imprecision given by the manufacturer is based on CLSI/National Committee for Clinical Laboratory Standard (NCCLS) EP5-A2: measurement of four samples of each concentration per day in two separate runs of two samples each for 20 days [4]. These imprecision results are lower than our findings indicating that imprecision based on the CLSI/NCCLS EP5-A2 protocol is not sufficient to provide a reliable $MD_{\text{cut-off}}$.

The $MD_{\text{cut-off } 7.0}$ were calculated from glucose concentration results obtained from three instruments run in parallel. The $MD_{\text{cut-off}}$ for individual instruments differed significantly in the long-term approach, but not in the short-term approach, indicating that a longer period of time is needed for an assessment of a reliable $MD_{\text{cut-off } 7.0}$. Furthermore, the $MD_{\text{cut-off}}$ for the individual instruments were slightly lower than those from the overall system; this result was expected since in addition to the larger imprecision, slight biases between the instruments also occurred, thereby adding to the variation when examining the overall distribution system 'instrument 4.'

The results reported to clinicians, however, represent the combined performance of all instruments. The results of this study show that even when each instrument is shown to have a

good assay performance, the clinician would experience differences in glucose results as large as 1.0 mmol/L at a glucose concentration of 6.6 mmol/L (minimum 6.2 mmol/L; maximum 7.2 mmol/L) due to imprecision and bias among the instruments. Therefore, the $MD_{\text{cut-off}}$ calculation should be based on MD distribution from all instruments that are connected to the automated distribution system.

Data from the long-term study show a reliable stability of the investigated systems over one decade, demonstrating the high quality of the manufacturer and medical laboratories. The effectiveness of the quality assurance systems, such as Rili-BAEK, with internal and external quality controls also contribute to the stable performance over time.

Short-term data showed minor performance shifts within the instruments, but these could not be linked to shifts in the workload. Even though the systems were observed to have a high stability, it has to be noted that the minimum QC frequency by Rili-BAEK allowed hundreds of patient results to be released in-between IQCs. An increase IQC frequency would reduce the number of released patient results, but it cannot be excluded that incidences between two IQC results may still occur unnoticed. Still, a higher IQC frequency reduces the number of samples that need to be retested after a failed IQC. It is up to the individual medical laboratory to carefully balance costs and quality of patient care.

The variability introduced by slight differences in performance within and between instruments can be covered by reporting the long-term $MD_{\text{cut-off}}$ across all connected instruments. Stable results for $MD_{\text{cut-off } 7.0}$ were obtained after about 30 independent control cycles of $MD_{\text{cut-off } 7.0}$, a number that was reached after 1 year of combining $MD_{\text{cut-off } 7.0}$ data from all three instruments.

The between-year $MD_{\text{cut-off } 7.0}$ values were not significantly different, except for the first year of the long-term approach when the instruments had just been introduced to the laboratory. Thus, medical laboratories may use the IQC of about 1 year to provide a reliable $MD_{\text{cut-off}}$ to report along with glucose concentration results used to diagnose DM. This

finding becomes especially important when close diagnostic cut-offs increase the need for reliable results to classify individuals correctly [15]. Subsequently, continuous monitoring of the $MD_{\text{cut-off}}$ builds an even more reliable database and can also facilitate identification of performance changes as well as comparisons of different measurement methods and laboratories.

Limitations

The study was limited to one type of instrument (Dimension Vista 1500), and the short-term imprecision part was limited to 1 week only. Adverse effects may have occurred less often than weekly and, therefore, these could not be identified in the short-term approach study due to its design. The commercially available IQC material used in the study closely resembles patient material, but is not patient material; therefore, effects due to any differences cannot be completely excluded. The aim of the study design was to assess measurement imprecision and express this imprecision as the MD.

CONCLUSION

Imprecision for glucose concentration measurements, assessed as $MD_{\text{cut-off } 7.0}$ from monthly IQC control cycles over a period of 10 years, was 0.44 mmol/L and, therefore, well below the recommended limit of 0.7 mmol/L. Hourly measurements made over a 1-week period confirmed these findings and illustrated the MD concept. Imprecision as measured by IQC is also remarkably stable over many years of operation.

Current imprecision assessment focuses only on single instruments, whereas clinicians perceive the combined analytical performance of all instruments used for a certain assay in a given laboratory. Therefore, we suggest deriving the $MD_{\text{cut-off}}$ from all instruments and control cycles that are used in the setting of patient care in a given medical laboratory. In our study, about 30 independent control cycles provided sufficient data to determine a reliable $MD_{\text{cut-off}}$. Establishing a continuous monitoring of $MD_{\text{cut-off}}$ may complement traditional quality assurance.

ACKNOWLEDGEMENTS

Funding. The study was in part funded by Bio-Rad. The Rapid Service Fee was paid for by the authors.

Medical Writing, Editorial, and other Assistance. The authors thank Stefan Pleus, Guido Freckmann and Erwin Schleicher for intense discussion of the findings and their helpful corrections of the manuscript.

Authorship. All named authors meet the International Committee of Medical Journal Editors (ICMJE) criteria for authorship for this article, take responsibility for the integrity of the work as a whole, and have given their approval for this version to be published.

Authorship Contributions. SK and AP drafted the manuscript; SZ, MD, AK and NF performed the data analyses; SK, AP, SZ, NF, AK and MN interpreted the data; NF, RL, LH, AK, MN revised the manuscript content. All authors read and approved the final manuscript. AP takes full responsibility for the work as a whole, including access to data and the decision to submit and publish the manuscript.

Disclosures. Sandra Keutmann, Stephanie Zylla, Mathilde Dahl, Nele Friedrich, Rüdiger Landgraf, Lutz Heinemann, Anders Kallner, Matthias Nauck and Astrid Petersmann have nothing to disclose. Astrid Petersmann is currently affiliated with the Institute of Clinical Chemistry and Laboratory Medicine, University Medicine Goettingen, Goettingen, Germany. Sandra Keutmann is currently affiliated with the Department of Anaesthesia, Trinity Hospital, Wesseling, Germany.

Compliance with Ethics Guidelines. This article does not contain any studies with human participants or animals performed by any of the authors.

Data Availability. The datasets generated and analyzed during the current study are

available from the corresponding author on reasonable request.

Open Access. This article is licensed under a Creative Commons Attribution-NonCommercial 4.0 International License, which permits any non-commercial use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

REFERENCES

1. Marathe PH, Gao HX, Close KL. American Diabetes Association standards of medical care in diabetes 2017. *J Diabetes*. 2017;9:320–4.
2. Müller-Wieland D, Petersmann A, Nauck M, et al. Definition, classification and diagnosis of diabetes mellitus. *Diabetol Stoffwechs*. 2016;11:S78–81.
3. World Health Organization. Use of glycated haemoglobin (HbA1c) in the diagnosis of diabetes mellitus: abbreviated report of a WHO consultation. WHO guidelines approved by the Guidelines Review Committee. Geneva: World Health Organization; 2011.
4. Joint Committee for Guides in Metrology (JCGM). Evaluation of measurement data—guide to the expression of uncertainty in measurement. JCGM 100:2008. <https://www.bipm.org/en/publications/guides/gum.html>.
5. Büttner J, Broth R, Broughton PM, Bowyer RC. International Federation of Clinical Chemistry. Committee on standards. Expert panel on nomenclature and principles of quality control in clinical chemistry. IFCC document stage 2, draft 3; 1979-07 with a view to an IFCC recommendation. Quality control in clinical chemistry. Part 4. Internal quality control. *J Clin Chem Clin Biochem* 1980;18:535–541.
6. Clinical and Laboratory Standards Institute (CLSI). Expression of measurement uncertainty in laboratory medicine; approved guideline. 2012:Document EP29.
7. Parvin CA. New insight into the comparative power of quality-control rules that use control observations within a single analytical run. *Clin Chem*. 1993;39:440–7.
8. Westgard JO, Barry PL, Hunt MR, Groth T. A multi-rule Shewhart chart for quality control in clinical chemistry. *Clin Chem*. 1981;27:493–501.
9. Revision of the “Guideline of the German Medical Association on quality assurance in medical lab examinations—Rili-BAEK” (unauthorized translation). *J Laboratory Med*. 2015;39(1):26–69.
10. Kunst A, Drager B, Ziegenhorn J. UV methods with hexokinase and glucose-6-phosphate dehydrogenase. In: Bergemeier H, editor. *Methods of enzymatic analysis*. Deerfield: Verlag Chemie; 1983. p. 163–72.
11. Badrick T, Cervinski M, Loh TP. A primer on patient-based quality control techniques. *Clin Biochem*. 2019;64:1–5.
12. Petersen PH, Ricos C, Stockl D, et al. Proposed guidelines for the internal quality control of analytical results in the medical laboratory. *Eur J Clin Chem Clin Biochem*. 1996;34:983–99.
13. Petersmann A, Wasner C, Nauck M, Kallner A. Frequency of extreme differences and clinical performance of glucose concentration measurements judged from 21,000 duplicate measurements. *Clin Chem*. 2013;59:998–1000.
14. Siemens Healthcare Diagnostics. Dimension Vista System reagent cartridge, glucose. Ref. K1039. Newark: Siemens Healthcare Diagnostics; 2017.
15. D’Orazio P, Burnett RW, Fogh-Andersen N, et al. Approved IFCC recommendation on reporting results for blood glucose (abbreviated). *Clin Chem*. 2005;51:1573–6.

14. Danksagung

Ich möchte Herrn Univ.-Prof. med. M. Nauck herzlich für die Überlassung der Arbeit im Institut für Klinische Chemie und Laboratoriumsmedizin danken.

Mein ganz besonderer Dank gilt der leitenden Oberärztin Frau PD Dr. med. Dipl.-Biol. A. Petersmann als meine Betreuerin, die mir immer mit Rat und Tat beiseite stand. Ohne Ihre Hilfe und Unterstützung wäre die Publikation, die Grundlage der Dissertation nicht zustande gekommen. Dies war keine Selbstverständlichkeit. Dafür bin ich sehr dankbar.

Allen Coautoren der wissenschaftlichen Arbeit, insbesondere bei der statistischen Auswertung Fr. S. Zylla, Fr. M. Dahl sowie Fr. N. Friedrichs möchte ich ganz herzlich für die Unterstützung danken.

Allen MTAs bin ich dankbar für Hilfe, durch die ich die praktische Ausführung im Labor bewerkstelligen konnte.

Frau Dipl.-Geogr. Sophia Lamp sowie Frau Dr. J. Schöpfel für ihre stets tatkräftige freundliche Hilfe und Unterstützung bei der Anfertigung der Arbeit.

Der IT/EDV für die Datenbereitstellung und Aufarbeitung im speziellen Hr. S. Bollmann sowie Hr. S. Westphal sowie allen übrigen Mitarbeitern des Institutes. Ebenfalls möchte ich Hr. T. Ittrich Service Ingenieur bei Siemens Healthcare Diagnostics GmbH danken für die Überlassung der Fehlermeldungen Vista Daten.

Ebenfalls möchte ich mich bei meiner Familie und meinen Freunden bedanken, welche die ganze Zeit über immer viel Verständnis für mich gehabt und mich unterstützt haben.