

Manuelle Medizin

Chirotherapie | Manuelle Therapie

Zeitschrift der Deutschen Gesellschaft für Manuelle Medizin

**Elektronischer Sonderdruck für
W.H. Harter**

Ein Service von Springer Medizin

Manuelle Medizin 2011 · 49:72–82 · DOI 10.1007/s00337-011-0823-1

© Springer-Verlag 2011

**zur nichtkommerziellen Nutzung auf der
privaten Homepage und Institutssite des Autors**

W.H. Harter

Was sind eigentlich Referenzdaten?

Überlegungen zur Beurteilung von biomechanischer Messung und
Referenzdatenkonzepten bei Patienten mit chronischen Rückenbeschwerden

www.ManuelleMedizin.springer.de

 Springer Medizin

Was sind eigentlich Referenzdaten?

Überlegungen zur Beurteilung von biomechanischer Messung und Referenzdatenkonzepten bei Patienten mit chronischen Rückenbeschwerden

Allgemeine Vorbemerkung – Notwendigkeit und Zweck

In der aktiven Behandlung von Patienten mit chronifizierten Rückenbeschwerden hat in den letzten Jahrzehnten zunehmend die Messung der Kraft der Wirbelsäulenstabilisierenden Muskulatur Eingang gefunden. Hier sind eine ganze Reihe von Therapie-/Geschäftskonzepten entstanden, welche zu therapeutischen Zwecken auf die systematische Beurteilung dieser Kraftmessungen basieren. Häufig werden die individuell erhobenen Daten mit anderen Messwerten als Stichprobenparametern, im Sinne eines Populationsvergleichs, verglichen und hieraus Therapieangebote entwickelt.

In dieser Arbeit soll – nicht zuletzt unter medizin-ethischen Gesichtspunkten – die Einbeziehung von ordentlichen Messungen/Messsystemen nach naturwissenschaftlichen Kriterien in die Diagnose (griechisch: „διάγνωση“, *diagnose*: Unterscheidung, Entscheidung, Erkenntnis) auf der Basis von Referenzdaten aus der medizinischen Trainingstherapie geklärt werden. Damit soll diese Arbeit einen Beitrag für Entscheidungsträger zur Beurteilung solcher therapeutischer Strategien im Gesundheitssystem leisten.

Sinn und Zweck eines ordentlichen Referenzdatenkonzepts soll die Entscheidungshilfe in der Diagnose sein. Ein individuelles spezifisches (gemessenes) Merkmal eines Patienten wird im Vergleich zu

einer repräsentativen Gesamtheit betrachtet. Mit diesem Vergleich und der annehmenden Ätiologie der Ursachen für eine gegebene individuelle krankheitsrelevante Abweichung sollen in der Beurteilung derselben Abweichung Schlüsse auf die Therapienotwendigkeit gezogen werden können. Dies hat weit reichende wissenschaftliche/medizinische und ökonomische Relevanz. Das bedeutet

- zum einen, dass der behandelnden Arzt valide Informationen aus Messungen zur Beurteilung einer gesundheitlichen Störung oder eines organischen Schadens erhält, diese in seine Differenzialdiagnostik einbezieht und hieraus Behandlungs-/Therapieempfehlungen unter nosologischen Gesichtspunkten ableitet,
- zum anderen, dass die Erfüllung des rekonditionierenden aktiven Therapieauftrags wieder ökonomische Konsequenzen (Kosten der Behandlung) für die Ressourcen des Gesundheitssystems nach sich zieht.

Kurze Betrachtung des Wesens naturwissenschaftlich begründeten Messens und medizinischer Relevanz

Die Kriterien einer Messung werden in der DIN 1319 1–4 geregelt. Sie regelt die

- Grundbegriffe von Messungen (1319/1),
- definiert die Messmittel in der Anwendung (1319/2) und

- den Umgang mit Messunsicherheiten/-fehlern und die Auswertung (1319/3/4).

Die Beobachtung eines Phänomens ist im Allgemeinen unvollständig, solang sie nicht zu einer quantitativen Information führt. [3]

In den Standardwerken der Physik [3] ist die Messung als Technik beschrieben, mit deren Hilfe eine physikalische Eigenschaft einer Zahl zugeordnet werden kann. Sie ist das Ergebnis eines Vergleiches mit einer standardisierten Größe. Die verfügbaren standardisierten normativen Größen basieren auf internationalen Konventionen im ISQ („International System of Quantities“) und beschreiben folgende Definitionen:

- Längen
- Masse
- Zeit
- Stromstärke
- Thermodynamische Temperatur
- Stoffmenge
- Lichtstärke

Aus diesen Basisgrößen werden, unter Kenntnis der Zusammenhänge, weitere Messgrößen hergeleitet. So sind die in der Biomechanik und Medizin – basierend auf den Modellen der klassischen (newtonschen) Mechanik – verwendeten Messgrößen wie Kraft, Energie, Leistung letztendlich Linearkombinationen der Basisgrößen. Aber auch die (Mess-)Angabe der Laborberichte basieren auf einer

Zusammenfassung · Abstract

Manuelle Medizin 2011 · 49:72–82 DOI 10.1007/s00337-011-0823-1
© Springer-Verlag 2011

W.H. Harter

Was sind eigentlich Referenzdaten? Überlegungen zur Beurteilung von biomechanischer Messung und Referenzdatenkonzepten bei Patienten mit chronischen Rückenbeschwerden

Zusammenfassung

Technische Systeme und biomechanische Messungen mit Hilfe von mehr oder weniger aufwändigen Analysesystemen sind auch aus der aktiven Behandlung von Patienten mit chronischen Rückenbeschwerden nicht mehr wegzudenken. Der Einsatz solcher biomechanischen Messungen der wirbelsäulenstabilisierenden Muskulatur als eine die Diagnose unterstützende Maßnahme ist in der Therapielandschaft zur mehr oder weniger regelmäßigen Übung geworden. Auch der Einsatz entsprechender apparativer Unterstützung in der aktiven Behandlung von Patienten nimmt zu. Die oftmals teure Anschaffung für die zur Messung geeigneter Analyse- und Therapiesysteme und die zunehmende Kostenübernahme der Analysen und der aktiven Therapiekonzepte rechtfertigen eine kritische

Auseinandersetzung mit der zugrundeliegenden Mess- und Interpretationssystematik.

Dabei bedienen sich unterschiedlichste Konzepte in der statistischen Interpretation der individuell gemessenen Werte eines Vergleichsdatenpools. Für die wenigsten hiervon ist die Herkunft der Vergleichsdaten durch Studien transparent gemacht worden. Ein solches Konzept – als repräsentativer Referenzdatenpool und die dazugehörige Datenbankentwicklung – lässt sich aber nur unter Einhaltung definierter internationaler Konventionen als statistisch belastbares und damit praktisch nutzbares darstellen.

In diesem Artikel werden die notwendigen Kriterien zur Messung des maximalen isometrischen Drehmoments als qualitativen biomechanischen Parameter des physio-

logischen und koordinativen Zustands der betroffenen Muskulatur als auch die Entwicklung von wertenden Referenzdaten zur Interpretation dieser Messungen aus naturwissenschaftlicher Sicht hergeleitet.

Es ergibt sich ein Kriterienkatalog, nach welchem die Güte von Referenzdaten und -konzepten von Behandlern und Kostenträgern beurteilt werden können. Praktische Beispiele zur Entwicklung und Pflege von statistisch belastbaren Datenpools erleichtern das Verständnis zu diesem komplexen Thema.

Schlüsselwörter

Biomechanik · Rückenbeschwerden · Referenzwerte · Datensammlung · Statistikmodell

What are reference data actually? Considerations on the assessment of biomechanical measurement and reference data in patients suffering from back pain

Abstract

Technical systems and biomechanical measurements with the help of biomechanical analysis systems are indispensable for active treatment of patients with chronic low back pain. The application of such biomechanical measurements of the back-stabilizing musculature as a measure supporting the diagnosis has become routine practice. The application of training devices in the active treatment of patients is also increasing. The acquisition of analysis and training devices is expensive and the increasing acceptance of the costs of the analyses and active therapy concepts justify a critical discussion of the underlying measuring and interpretation systems.

The very different concepts are useful for statistical interpretation of the individual measured values with the help of comparative data pools. For a few the origin of the comparative data pools has been made clear by controlled studies. However, such a concept with reference values and a data base can only be demonstrated as being statistically sound under observance of defined international conventions and therefore of practical use.

In this article the criteria for the necessity and measurement of the maximum isometric torque as a qualitative biomechanical parameter for the physiological and coordinative

state of the affected musculature are derived. The development of reference values for the interpretation of these measurements is also shown from a scientific viewpoint.

A catalogue of criteria for assessment of the benefit of reference values and their use is constructed. Practical examples for the development and maintenance of a statistically sound data base facilitate an understanding of this complicated subject.

Keywords

Biomechanics · Low back pain · Reference values · Data collection · Statistical model

Übersichten

(modellhafte) Interpretation der Basis-einheiten.

Streng genommen wird aus der Kenntnis physikalischer, chemischer und biologischer Zusammenhänge (Modelle) noch zwischen

- direkten (Messung aufgrund eines Vergleichs mit einem vorher definierten Maßstabs) und
- indirekten Messungen (Messung einer bekannten Größe, die zu der zu messenden Größe in einem bekannten Zusammenhang steht)

unterschieden. Dies hat ganz besondere Bedeutung in der Interpretation des quantitativen Ergebnisses. Bei der direkten Messung ist die Interpretation trivial. Misst man die Länge eines Raumkörpers durch Anlegen eines Zollstocks, so vergleicht man ein und dieselbe Messgröße des zu messenden Objekts mit dem normierten Maßstab. *Komplizierter* wird dies schon am trivialen Beispiel der Interpretation bei der Fiebermessung mit einem veralteten Quecksilberthermometer sichtbar. Direkt gemessen wird hier nicht die Körpertemperatur, sondern die Ausdehnung der Quecksilbersäule in der Glashöhle. Aus der Kenntnis der thermodynamischen Zusammenhänge schließen wir auf die Umgebungstemperatur (z. B. Mundhöhle), in welcher gemessen wurde und welche die Ausdehnung verursacht hat. Mit der Feststellung/Ableitung des Wertes ist der Messvorgang beendet. Jetzt beginnt erst der eigentliche Interpretationsvorgang:

Der Messwert der individuellen Körpertemperatur wird mit der Norm (MW \pm Standardabweichung = $36,75 \pm 0,75^\circ\text{C}$) verglichen.

Eine Abweichung kann entweder durch äußere Einflüsse (Kälte, Unterkühlung, Hitze) oder durch eine gezielte Regulation der Sollwertverstellung im Zwischenhirn aus Krankheitsgründen (Infekt) verursacht sein.

Erst wenn diese – unter Einbeziehung der Umgebungsvariablen und insbesondere der Anamnese (Beispiel: Husten, Schnupfen, Heiserkeit) und weiterer (Krankheits-)Merkmale – abgeschlossen ist, führt dies zur Erkenntnis und Entscheidung (Diagnose s. oben). Mit dieser Interpretation ist dann auch im öko-

nomischen und juristischen Sinn die Verantwortung, insbesondere für das daraus folgende Handeln, verbunden.

Die institutionelle Verantwortung für die Interpretation von Messungen, selbst bei einem wie oben beschriebenen trivialen Beispiel und die daraus folgenden Konsequenzen

- Therapieentscheidung (Medikamentenverordnung, Operationsindikation etc.)
- „Auslösen des Versicherungsfalls“, Krankenversicherung
- Krankschreibung (Lohnentgeltfortzahlung, Krankengeldbezug)

u. v. a. ist prinzipiell und ausschließlich dem approbierten Arzt vorbehalten! Einerseits ist eine Messung der Körpertemperatur ein *Standardverfahren*, eine regelmäßige häusliche Übung, andererseits kann eine Fehlentscheidung auch mit weitreichenden Konsequenzen und Schäden verbunden sein. Diese Konsequenzen müssen aus ethischen und ökonomischen (Haftung) Gründen geregelt sein.

Beurteilung konditioneller und koordinativer Fähigkeiten

Messung der Kraft und Interpretation des funktionellen Zustands der beteiligten Muskulatur

Im einfachsten Fall schließen wir von der *Art und Weise* mit der z. B. ein Gewicht von einem Gewichtheber gehoben wird unwillkürlich auf die Kraft der handelnden Person und damit auch auf den (funktionellen/strukturellen) Zustand seiner Muskulatur. Letztendlich ist dies auch eine Messung, da – unter den Bedingungen eines *standardisierten Verfahrens*, der sportartbedingten technischen Regelvorschrift – der Beobachtung eindeutig eine Zahl (das gehobene Gewicht) zugeordnet werden kann. Es handelt sich demnach auch um eine direkte Messung, da die Kraft, welche vom Sportler erzeugt wird, im Sinne des dritten newtonschen Axioms exakt der Gegenkraft entspricht, mit welcher das zu hebende Gewicht durch die Gravitation angezogen wird.

Hieraus aber auf den organischen Zustand und damit auf die Gesundheit/

Krankheit des/der beteiligten Muskeln zu schließen, nämlich die

- Struktur (Faserverteilung, Qualität und Quantität der der muskulären Zellstruktur etc.)
- Funktion (neuromuskuläre Ansteuerung, intra- und intermuskuläre Inervation),

wäre höchst fahrlässig. Letztendlich ist die Bewegungsausführung von einer großen Zahl determinierender Faktoren (konditionelle, koordinative, kognitive und afektive Voraussetzungen) abhängig.

Zur Beurteilung der im therapeutischen Markt angebotenen Kraftmessmethoden muss daher äußerst kritisch hinterfragt werden:

- a) Welches Phänomen soll beobachtet werden?
- b) Bildet die eingesetzte Messmethode das zu beobachten gewünschte Phänomen testerunabhängig homomorph und damit objektiv ab (Testgütekriterien: Validität, Objektivität, Reliabilität)?
- c) Ist das Messsystem unter ingenieurwissenschaftlichen Gründen tauglich (Messgenauigkeit)?
- d) Unter welchen Voraussetzungen (Zuweisungsvorschrift) ist der zu messende numerische Wert dem zu beobachten gewünschten Phänomen zuzuordnen?
- e) Welche Zusammenhänge (Kognitionen, weitere physische und psychische Einflussfaktoren) erzeugen Varianten, die streng genommen größer sind als die messtechnische Genauigkeit des Systems?

Das bedeutet nicht nur, dass der Bewegungsausführung in einem Analysegerät auf einer Anzeige (Computermonitor) ein Wert zugeordnet wird, welcher als Kraft oder als Drehmoment technisch interpretierbar ist, sondern auch, dass eine geeignete Zuweisungsvorschrift existiert, nach welcher der Messwert interpretierbar ist. Dabei sind die Punkte a) d) und e) über wissenschaftliche Studien und Überlegungen darzulegen. Auf diese wird im folgenden Kapitel eingegangen. Die Punkte b) und c) sollten – im Rahmen des Medizinproduktegesetzes (MPG, nationale Umsetzung der Regel 12 europäischen

Richtlinien 90/42/EWG) – durch den Hersteller verargumentiert werden und die technische Güte der Herstellung und der Messtechnik dokumentieren.

Dieser Einschätzung wird endgültig dadurch Rechnung getragen, dass diese Kriterien durch eine klinische Prüfung im Rahmen des MPG werksseitig nachzuweisen sind! Nicht nur aus Gründen der Qualitätssicherung, sondern auch aus der Sicht des ärztlichen Berufsethos (Einbeziehung von Messungen in die Differenzialdiagnostik) ist der klinische Nachweis dieser Kriterien überprüfbar sicherzustellen.

(1) Die Eignung von Medizinprodukten für den vorgesehenen Verwendungszweck ist durch eine klinische Bewertung anhand von klinischen Daten nach § 3 Nummer 25 zu belegen, soweit nicht in begründeten Ausnahmefällen andere Daten ausreichend sind. Die klinische Bewertung schließt die Beurteilung von unerwünschten Wirkungen sowie die Annehmbarkeit des in den Grundlegenden Anforderungen der Richtlinien 90/385/EWG und 93/42/EWG genannten Nutzen-/Risiko-Verhältnisses ein. Die klinische Bewertung muss gemäß einem definierten und methodisch einwandfreien Verfahren erfolgen und gegebenenfalls einschlägige harmonisierte Normen berücksichtigen. (§ 19 MPG Klinische Bewertung, Leistungsbewertung)

Die Hersteller bieten die Geräte i. d. R. als Mess- und Analysesysteme an. Nach prinzipiellen Überlegungen zum MPG, aber auch zum Wettbewerbsrecht, sollte der Hersteller hierzu nicht nur die Anweisungen zum sinnvollen Gebrauch der Systeme, sondern auch zur standardisierten Messung und in diesem Zusammenhang auch den Nachweis der Testgütekriterien vorhalten (*unerwünschte Wirkung* durch falsche Messverfahren). Sollte dann ein Anwender, ggf. zu Studienzwecken oder aus anderen Erwägungen, von diesen Vorgabe abweichen wollen, so muss er selbsttätig den wissenschaftlichen Nachweis der Objektivität, Reliabilität und Validität unter den von ihm gewünschten Standardisierungsbedingungen sicherstellen.

Interpretation des Messergebnisses

Das erste qualitative und quantitative Beurteilungssystem, welches in dieser Hinsicht angeboten wurde, ist die Muskel-

funktionsdiagnostik nach Janda [14]. Sie hat bis heute ihre Tauglichkeit und Praktikabilität – trotz aller hochaufwändigen technischen Systeme – nicht eingebüßt. Analog zum naturwissenschaftlich geforderten Informationsgewinn hat Janda den zu beurteilenden Menschen mit seinen erforderlichen funktionellen Fähigkeiten der Einfachheit halber zum eigenständigen Messsystem erklärt. Die zentrale Frage seiner qualitativen Beurteilung ist:

- Kann der zu beurteilende Patient das für sein *System Körper* zu definierende, nämlich alltagstaugliche Teilkörperfewicht (Arm, Kopf, Oberkörper) mit welcher eindeutig abgrenzbaren Qualität bewegen?
 1. Die Muskulatur ist imstande, einen der Bewegung von außen entgegengesetzten Widerstand zu überwinden.
 2. Die Muskulatur kann nur noch die Schwerkraft überwinden.
 3. Die Muskulatur kann Körperteile nur noch unter Ausschluss der Schwerkraft überwinden.
 4. Es kommt nur eine Muskelanspannung zustande, eine Bewegung bleibt dagegen aus.

Allerdings sind hierbei kritisch die Grenzen bei der qualitativen *Stufe 1* gesetzt. Die Kriterien des Muskelfunktionstests nach Janda grenzen in den Kategorien 2–4 eine schwergradige funktionelle/strukturelle Morbidität von der Alltagstauglichkeit ab. Die Beurteilung der tatsächlichen Alltagstauglichkeit ist hierbei quantitativ zur Normalität nur schlecht weiter differenzierbar. Insbesondere, da auch bei Patienten, welche bei Janda der *Stufe 1* entsprechen, trotzdem sowohl schwere Rückenbeschwerden vorliegen können als auch die Muskulatur funktionell und strukturell schon massiv beeinträchtigt im pathologischen Sinne sein kann.

Das pathogenetische Prinzip der wirbelsäulenstabilisierenden Muskulatur bei chronischen Rückenbeschwerden

Funktionelle, insbesondere muskuläre, aber auch koordinative, defizitäre Zustände beeinträchtigen den Patienten in der individuellen Handlungskompetenz

Hier steht eine Anzeige.

 Springer

zur Erfüllung seiner alltäglichen Aufgaben. Die individuellen Eigenschaften und Fähigkeiten (Kapazität) eines Menschen bestimmen die Erfüllung der objektiven Aufgaben und Belastung.

Wenn diese hier funktionelle Kapazität der Alltagstauglichkeit, insbesondere im Vergleich mit einer gesunden Population

- individuell reduziert ist,
- ein kausaler Zusammenhang mit einem Krankheitsbild besteht und
- die funktionelle Störung das Risiko, dass eine bestehende Erkrankung wieder auftritt oder sich verschlechtert, nachweislich erhöht ist,

so ist die Voraussetzung für ein therapeutisches Vorgehen erfüllt. Die *reduzierte funktionelle Kapazität* entspricht dann einem Risikofaktor [13].

Eine Maßnahme, welche die Reduktion eines Risikofaktors einer schon bestehenden Erkrankung bewirkt, entspricht demnach einer kausalen Therapie und bedingt nach der ärztlichen Diagnose die Einleitung kurativer und tertiärpräventiver Maßnahmen.

Die funktionelle Störung

Patienten mit chronischen Rückenbeschwerden haben nachweislich signifikant häufiger muskuläre Defizite gegenüber untrainierten Normalpersonen [4, 9, 10, 11, 16, 17, 20, 21, 22]. Die Notwendigkeit der muskulären Rekonditionierung bei Patienten mit chronischen Rückenbeschwerden wird in der Literatur vielfach beschrieben und begründet. Harter [8] hat die Zusammenhänge von chronischen Rückenbeschwerden und messbaren muskulären Defiziten als organischen und somit behandlungsbedürftigen Schaden beschrieben.

Eine weitergehende funktionelle Beurteilung und Differenzierung in betroffene Muskelgruppen und muskuläre Dysbalancen sollte durch die Messung der isometrischen Maximalkraft, idealerweise ergänzt durch die Verfügbarkeit von Referenzdaten nach internationalen Konventionen (vgl. International Federation of Clinical Chemistry, IFCC bei Solberg [23, 24]), abgesichert werden.

Harter [8] hat beschrieben, wie eine besonders signifikante individuelle Abweichung von einer repräsentativen Norma-

lität zu werten ist. Demnach muss der behandelnde Arzt bei einer entsprechenden Anamnese und Chronifizierung der Beschwerden in Verbindung mit langanhaltenden Rezidiven regelmäßig von auffälligen und damit behandlungsbedürftigen strukturellen und funktionellen Veränderungen der betroffenen Muskulatur ausgehen. Uhlig [25] hat den medizinischen Zusammenhang in der Differenzialdiagnostik exakt beschrieben. Die messtechnische Dokumentation des (pathologischen) Funktionszustand der wirbelsäulenstabilisierenden Muskulatur und deren Beseitigung durch aktivierende Therapiemaßnahmen [8] findet damit seinen Platz innerhalb der komplexen Differenzialdiagnostik [25] entsprechend der europäischen Leitlinien [2]. Auch wenn von interessierter Seite die aktive Therapie häufig monokausal formuliert wird, ist eine schlüssige Therapierung des komplexen biopsychosozialen Geschehens nur innerhalb eines komplexen therapeutischen Curriculums mit einer sowohl funktionell-somatischen und didaktisch-psychologischen, ggf. sozialtherapeutischen Intervention zu erwarten.

Demnach ist – eingedenk weiterer differenzialdiagnostischer Kriterien – bei relativ kurzen Episoden der Einschränkung und/oder Immobilität der wirbelsäulenstabilisierenden Muskulatur schon von signifikanten und damit behandlungsbedürftigen funktionellen und strukturellen Veränderungen des betroffenen Bewegungsorgans auszugehen, welche nachweislich i. d. R. nicht mehr durch den Alltag kompensiert wird. Damit hat der Patient auch nicht mehr das funktionelle/strukturelle organische Potenzial, das ihn im ursprünglich gesunden Zustand in seiner funktionellen Kapazität für seine Alltagstauglichkeit repräsentierte. Eine individuelle Abweichung des damit für den betroffenen Patienten möglichen maximalen isometrischen Drehmoments von einer alters- und geschlechtsentsprechenden repräsentativen Stichprobe dokumentiert dabei – nichtinvasiv und in der Durchführung vergleichsweise kostenextensiv – den organischen Schaden. Damit ist aber auch mit derart präzisen Messungen die *diagnostische Lücke* zwischen der oben beschriebenen Muskelfunktionsdiagnostik nach Janda bei Patienten mit einer funktionellen Graduierung von 1 zur beschwer-

defreien Normalperson nicht nur qualitativ, sondern auch quantitativ geschlossen.

Was sind Referenzdaten?

Die Entwicklung von Referenzdaten zu einem definierten Merkmal hat inzwischen in den unterschiedlichsten Disziplinen Eingang gefunden. So werden Referenzdaten als Kontrollinstrument in der Betriebswirtschaft genauso eingesetzt wie in der chemischen Industrie, insbesondere aber in der Biochemie. Hier ist es v. a. die *International Federation of Clinical Chemistry* (IFCC), welche zur Entwicklung und Verwendung von Referenzdaten internationale verbindliche Standards gesetzt hat.

Die Vorstellungen darüber, welche Daten zur Referenz für ein zu beschreibendes Merkmal herangezogen und/oder entwickelt werden, divergieren allerdings erheblich. Dies geht teilweise über die willkürliche Bestimmung von *Erfahrungswerten* bis hin zur exakten stochastischen Ermittlung eines statistischen Klons, mit dem die repräsentative Wirklichkeit eines Phänomens auch entscheidungsunterstützend abgebildet werden soll.

In dem besonderen Fall von Patienten mit chronischen Rückenschmerzen bedeutet dies, dieselben mit einer repräsentativen Menge beschwerdefreier Menschen zu vergleichen, um ursächlich auf die funktionelle und/oder strukturelle Störung schließen zu können.

Definition der Referenzwertverteilung und Pflichtenkatalog

Eine Referenzstichprobe sollte eine *repräsentative Abbildung* der Wirklichkeit, in diesem Falle der gesundheitlich bezüglich Rückenschmerzen unbeeinträchtigten Normalbevölkerung, sein. Oftmals werden erhobene Stichproben präsentiert, welche allein von der Größe her hierzu statistisch nicht belastbar sind.

Als Beispiel zur Abschätzung relevanter Stichprobengrößen soll hier die Entwicklung der Fallzahlen für eine A-priori-Stichprobe (Erfassung einer geeigneten Zufallsstichprobe) als statistisch belastbarer Idealfall vorgestellt werden. Als weitere Möglichkeit gibt Solberg [24] die Entwicklung einer A-posterior-Stichprobe (Auswahl geeigneter Referenzpersonen aus einem be-

stehenden Datenpool nach beschriebenen Kriterien) an. Ein Beispiel hierzu mit entsprechenden Fallzahlen wird nach dieser Entwicklung beschrieben.

Folgende Begriffe müssen innerhalb des Designs der Referenzdatenentwicklung abgeklärt und überprüfbar sein:

- Eine *Referenzpopulation* umfasst Personen, die hinsichtlich des zu messenden Merkmals gesund und unbelastet sind.
- *Referenzwerte einer Messmethode* sind an einer repräsentativen Stichprobe aus der Referenzpopulation unter definierten Bedingungen gewonnene Werte, die die natürliche, normale biologische Variation darstellen.
- *Referenzgrenzen* beschreiben ein Intervall, dessen Überdeckung 95% der Referenzwerte beträgt.

Diese Definitionen ergeben sich insbesondere aus der Literatur um die Referenzwertentwicklung und den Richtlinien für Referenzwerte ([5, 12, 18, 19, 23, 24, 26]).

Hieraus ergeben sich zur Entwicklung folgende Aufgaben für die Durchführung:

- a) Messmethode definieren,
- b) Referenzpopulation definieren,
- c) repräsentative Stichprobe ziehen,
- d) unter Qualitätskontrolle messen,
- e) Referenzwerte darstellen,
- f) Referenzgrenzen bestimmen,
- g) stochastisches Rauschen reduzieren.

Hierzu sind folgernde Vorbereitungen zu treffen:

Messmethode definieren und Sicherstellung der inhaltlichen Qualitätskontrolle (a, d). Die Messmethode muss strukturiert und standardisiert werden. Das standardisierte Protokoll dient als Vorlage und ist zur intensiven Beschulung teilnehmender Institute/Tester nachzuweisen[8].

Referenzpopulation definieren und repräsentative Stichprobe ziehen/Kalkulation der notwendigen Fallzahlen (b, c). Die nachzuprüfenden Fallzahlen für eine A-priori-Stichprobe könnten sich aus einer Abschätzung nach den Erwartungswerten der 12. bereinigten Bevölkerungsvorausberechnung des Bundesamtes für Statistik (www.destatis.de/bevoelkerungs-pyramide/) für das Jahr 2010 ergeben.

Tab. 1 Erwartete Bevölkerungszahlen in der Bundesrepublik 2010 nach Alter und Geschlecht

Alter	Mio.	%	m	m %	w	w %
<20 Jahre	15,3	19	7,498	9,31	7,802	9,69
20–29 Jahre	9,9	12	4,851	5,88	5,049	6,12
30–39 Jahre	10	12	4,900	5,88	5,100	6,12
40–49 Jahre	13,9	17	6,812	8,33	7,088	8,67
50–59 Jahre	11,5	14	5,635	6,86	5,865	7,14
>59 Jahre	21,2	26	10,389	12,74	10,811	13,26
Gesamt	81,7		40,085		41,715	

m (%), w (%) prozentualer Anteil Männer/Frauen; m Fallzahl Männer in Mio., w Fallzahl Frauen in Mio.

Tab. 2 Minimale Fallzahl bei t-Verteilung

Alter	m	w
<20 Jahre	54	54
20–29 Jahre	35	35
30–39 Jahre	35	35
40–49 Jahre	49	49
50–59 Jahre	41	41
>59 Jahre	75	75
Gesamt	288	290
Relativfaktor zur kleinsten Gruppe	138,6	144,3

Tab. 3 Minimale Fallzahl bei Normalverteilung

Alter	m	w
<20 Jahre	187	186
20–29 Jahre	121	120
30–39 Jahre	123	121
40–49 Jahre	170	169
50–59 Jahre	141	140
>59 Jahre	260	257
Gesamt	1002	993
Relativfaktor zur kleinsten Gruppe	40,425	42,075

Entwicklung einer A-priori-Stichprobe

Sinnvollerweise sollen die Gesamtstichproben nach Geschlecht und Alter gruppenweise stratifiziert werden. Dabei werden Männer und Frauen jeweils nach Dekaden aufgeteilt. Damit dies auch repräsentativ entwickelt werden kann, sollen die tatsächlichen und die zu erwartenden Zahlen für das Bemessungsjahr 2010 aus der 12. bereinigten Bevölkerungsvorausberechnung entwickelt werden.

Aus der dekadensweise entwickelten Verteilung der Erwartungswerte für das Jahr 2010 wurde der Proporz alters- und geschlechtsspezifisch errechnet. Der Proporz berechnete sich nach den tatsächlichen Fallzahlen der Geschlechterverteilung aus dem Jahr 2008.

Laut dem Bundesamt für Statistik betrug in diesem Jahr die Anzahl der Gesamtbevölkerung N=82,0024 Mio. Einwohner. Bei einer Fallzahl von n=40,1843 Mio. männlichen und n=41,8181 Mio. weiblichen Mitbürgern ergeben sich jeweils der relative Proporzfaktor von 2,04 (Männer) und 1,96 (Frauen) welche der Berechnung in **Tab. 1** zugrundeliegen ([www.destatis.de/jetspeed/portal/cms/Sites/destatis/Internet/DE/Content/Statistiken/Bevoel](http://www.destatis.de/jetspeed/portal/cms/Sites/destatis/Internet/DE/Content/Statistiken/Bevoelkerungs-pyramide/)

kerung/Bevoelkerungsstand/Tabellen/Content50/GeschlechtStaatsangehoerigkeit.psm).

Es werden 2 Modelle gerechnet:

- Modell 1: Stratifikationsgruppen entsprechender t-Verteilung (geringere statistische Belastbarkeit)
- Modell 2: Stratifikationsgruppen entsprechen mindestens der z-Verteilung (hohe statistische Belastbarkeit)

Bei einer notwendigen Minimalanzahl von n=35 für eine t-Verteilung – als minimale statistische belastbare Fallzahl in einer Zufallsstichprobe zur Berechnung von Verteilungsmerkmalen (vgl. Solberg [23]) – angewendet auf die jeweils kleinste Stratifikationsgruppe (getrennt nach Geschlecht) aus **Tab. 1** ergeben sich die Fallzahlen für das Modell 1 zur Fallzahlkalkulation (**Tab. 2**):

Sollen die Studien mit deutlich höherer statistischer Belastbarkeit entwickelt werden, so ist – unter Voraussetzung der Normalverteilung in einer Zufallsstichprobe in der jeweils kleinsten Stratifikationsgruppe (getrennt nach Geschlecht) – eine minimale Fallzahl in der kleinsten Gruppe von n=120 vorauszusetzen. Damit ergibt sich der jeweilige Proporz für das Modell 2 in **Tab. 3**.

Übersichten

Tab. 4 Referenzwerte und Varianzanalyse der jährlichen Veränderung im Bemessungszeitraum (Alter 10–19 Jahre)

	Frauen (n)	MW ± SD (Nm)	95% CI (Nm)	OWA (sign.)	Änd./Jahr	Männer (n)	MW ± SD (Nm)	95% CI (Nm)	OWA (sign.)	Änd./Jahr
Extension	175	4,01±1,27	4,23–1,40	0,042	–5%	257	5,03±1,62	4,83–5,23	0,337	–5%
Flexion	167	2,66±0,66	2,56–2,76	0,631	11%	254	3,28±0,86	3,18–3,39	0,020	–9%
Rotation rechts	171	1,96±0,51	1,88–2,04	0,979	0%	250	2,69±0,79	2,60–2,79	>0,001	–18%
Rotation links	171	1,93±0,54	1,85–2,01	0,668	–3%	250	2,71±0,77	2,61–2,80	>0,001	–15%
Lateralflexion rechts	160	2,31±0,61	2,21–2,41	0,836	–5%	248	2,93±0,89	2,82–3,04	>0,001	–19%
Lateralflexion links	160	2,36±0,65	2,00–4,29	0,582	–10%	248	2,97±0,82	2,87–3,08	>0,001	–19%

MW Mittelwert, SD Standardabweichung, 95% CI 95% Konfidenzintervall, OWA Ergebnis der Varianzanalyse über die Änderung der jährlichen Werte im Bemessungszeitraum, Änd./Jahr durchschnittliche Änderung des Mittelwerts pro Jahr.

Tab. 5 Referenzwerte und Varianzanalyse der jährlichen Veränderung im Bemessungszeitraum (Alter 20–29 Jahre)

	Frauen (n)	MW ± SD (Nm)	95% CI (Nm)	OWA (p)	Änd./Jahr	Männer (n)	MW ± SD (Nm)	95% CI (Nm)	OWA (p)	Änd./Jahr
Extension	574	4,35±1,35	4,24–4,46	0,009	8%	523	5,89±1,48	5,76–6,02	0,010	–12%
Flexion	550	2,75±0,81	2,68–2,81	0,001	9%	515	3,70±0,88	3,62–3,77	0,140	1%
Rotation rechts	546	2,00±0,58	1,96–2,05	0,311	–3%	495	3,32±0,83	3,25–3,39	>0,001	–5%
Rotation links	546	2,03±0,61	1,98–2,08	0,003	–3%	493	3,37±0,85	3,29–3,44	>0,001	–7%
Lateralflexion rechts	510	2,49±0,67	2,43–2,55	>0,001	–6%	483	3,59±0,87	3,52–3,67	>0,001	–13%
Lateralflexion links	509	2,54±0,66	2,48–2,59	0,001	–3%	485	3,61±0,84	3,54–3,69	>0,001	–12%

MW Mittelwert, SD Standardabweichung, 95% CI 95% Konfidenzintervall, OWA Ergebnis der Varianzanalyse über die Änderung der jährlichen Werte im Bemessungszeitraum, Änd./Jahr durchschnittliche Änderung des Mittelwerts pro Jahr.

Tab. 6 Referenzwerte und Varianzanalyse der jährlichen Veränderung im Bemessungszeitraum (Alter 30–39 Jahre)

	Frauen (n)	MW ± SD (Nm)	95% CI (Nm)	OWA (p)	Änd./Jahr	Männer (n)	MW ± SD (Nm)	95% CI (Nm)	OWA (p)	Änd./Jahr
Extension	984	4,24±1,25	4,17–4,32	0,003	5%	1273	5,54±1,49	5,46–5,62	0,000	–10%
Flexion	920	2,67±0,73	2,63–2,72	0,006	1%	1232	3,40±0,84	3,36–3,45	0,000	0%
Rotation rechts	922	2,00±0,59	1,97–2,04	0,036	–3%	1218	3,06±0,78	3,02–3,11	0,030	–3%
Rotation links	921	2,03±0,62	1,99–2,07	0,004	–1%	1218	3,12±0,79	3,08–3,17	>0,001	–5%
Lateralflexion rechts	807	2,35±0,67	2,30–2,39	>0,001	–3%	1160	3,16±0,83	3,11–3,21	>0,001	–17%
Lateralflexion links	807	2,45±0,69	2,40–2,49	>0,001	0%	1161	3,25±0,80	3,20–3,29	>0,001	–12%

MW Mittelwert, SD Standardabweichung, 95% CI 95% Konfidenzintervall, OWA Ergebnis der Varianzanalyse über die Änderung der jährlichen Werte im Bemessungszeitraum, Änd./Jahr durchschnittliche Änderung des Mittelwerts pro Jahr.

Tab. 7 Referenzwerte und Varianzanalyse der jährlichen Veränderung im Bemessungszeitraum (Alter 40–49 Jahre)

	Frauen (n)	MW ± SD (Nm)	95% CI (Nm)	OWA (p)	Änd./Jahr	Männer (n)	MW ± SD (Nm)	95% CI (Nm)	OWA (p)	Änd./Jahr
Extension	1077	4,20±1,29	4,13–4,28	0,930	–4%	1190	5,52±1,43	5,44–5,60	0,070	–4%
Flexion	992	2,64±0,74	2,59–2,68	0,030	–1%	1159	3,30±0,78	3,25–3,34	0,030	0%
Rotation rechts	999	1,97±0,62	1,93–2,01	>0,001	–4%	1147	3,01±0,73	2,97–3,05	>0,001	–1%
Rotation links	998	2,00±0,63	1,96–2,04	>0,001	–3%	1147	3,10±0,75	3,06–3,15	>0,001	–5%
Lateralflexion rechts	856	2,26±0,68	2,22–2,31	>0,001	–5%	1082	3,03±0,73	2,99–3,07	>0,001	–8%
Lateralflexion links	856	2,36±0,70	2,32–2,41	>0,001	–2%	1082	3,09±0,76	3,05–3,14	>0,001	–12%

MW Mittelwert, SD Standardabweichung, 95% CI 95% Konfidenzintervall, OWA Ergebnis der Varianzanalyse über die Änderung der jährlichen Werte im Bemessungszeitraum, Änd./Jahr durchschnittliche Änderung des Mittelwerts pro Jahr.

Die allgemeinen Merkmale können sich nach dem Körpermormkonzept nach Israel [13] definieren. Damit wird das relative Normkonzept zur Abgrenzung einer gesunden von einer kranken Population sichergestellt. Inhaltlich bedeutet dies:

- momentane Rückenbeschwerden ja/nein (heute und/oder die letzten 7 Tage)

- bei gleichzeitiger Dauer der Beschwerde in Jahren (<1 Jahr);
- bei Rücken-/Nackenbeschwerden jeweils „ja“;
- Dauer aktueller Beschwerden in Wochen (<6 Wochen);
- Schmerzintensität (Rücken/Nacken; visuelle Analogskala keine Schmerzen

0–10 unerträgliche Schmerzen mindesten <5);

- Rezidivhäufigkeit unregelmäßig.

Probanden, die dieser Merkmalskombination entsprechen, können in die Referenzpopulation einbezogen werden.

Tab. 8 Referenzwerte und Varianzanalyse der jährlichen Veränderung im Bemessungszeitraum (Alter 50–59 Jahre)

	Frauen (n)	MW ± SD (Nm)	95% CI (Nm)	OWA (p)	Änd./ Jahr	Männer (n)	MW ± SD (Nm)	95% CI (Nm)	OWA (p)	Änd./ Jahr
Extension	635	4,08±1,26	3,98–4,18	0,298	1%	787	5,40±1,42	5,30–5,50	0,010	–8%
Flexion	593	2,50±0,70	2,45–2,56	0,900	–3%	751	3,19±0,73	3,13–3,24	0,020	0%
Rotation rechts	587	1,89±0,58	1,85–1,94	0,340	–1%	741	2,98±0,72	2,93–3,03	0,020	–3%
Rotation links	587	1,93±0,61	1,88–1,98	0,030	–4%	741	3,06±0,73	3,01–3,11	0,010	–3%
Lateralflexion rechts	513	2,15±0,63	2,10–2,20	>0,001	–2%	687	2,93±0,71	2,88–2,98	0,020	–3%
Lateralflexion links	513	2,25±0,63	2,20–2,31	0,030	0%	687	3,03±0,72	2,97–3,08	0,030	–3%

MW Mittelwert, SD Standardabweichung, 95% CI 95% Konfidenzintervall, OWA Ergebnis der Varianzanalyse über die Änderung der jährlichen Werte im Bemessungszeitraum, Änd./Jahr durchschnittliche Änderung des Mittelwerts pro Jahr.

Tab. 9 Referenzwerte und Varianzanalyse der jährlichen Veränderung im Bemessungszeitraum (Alter >59 Jahre)

	Frauen (n)	MW ± SD (Nm)	95% CI (Nm)	OWA (p)	Änd./ Jahr	Männer (n)	MW ± SD (Nm)	95% CI (Nm)	OWA (p)	Änd./ Jahr
Extension	243	3,83±1,19	3,68–3,98	0,768	1%	432	5,11±1,36	4,98–5,23	0,290	1%
Flexion	218	2,24±0,70	2,15–2,33	0,280	11%	411	2,98±0,74	2,91–3,05	0,060	–4%
Rotation rechts	231	1,75±0,66	1,67–1,84	0,138	3%	419	2,79±0,72	2,72–2,86	0,020	6%
Rotation links	230	1,76±0,64	1,68–1,85	0,211	5%	419	2,83±0,74	2,76–2,91	0,040	5%
Lateralflexion rechts	184	1,92±0,57	1,84–2,01	0,317	3%	382	2,69±0,72	2,62–2,77	0,060	–5%
Lateralflexion links	186	2,06±0,61	1,97–2,14	0,617	–4%	382	2,77±0,70	2,70–2,84	0,100	–3%

MW Mittelwert, SD Standardabweichung, 95% CI 95% Konfidenzintervall, OWA Ergebnis der Varianzanalyse über die Änderung der jährlichen Werte im Bemessungszeitraum, Änd./Jahr durchschnittliche Änderung des Mittelwerts pro Jahr.

Durchführungskriterien zur Messung, Datenschutz und Datensammlung

Jeder Proband, welcher der Referenzpopulation nach dem Körpernormkonzept entspricht (Befragung auf Formblatt) sollte eine Datenerklärung unterschrieben haben, dass er mit der anonymisierten Weiterverwendung seiner Daten einverstanden ist. Diese Erklärung sollte im Original der studienleitenden Stelle postalisch übersendet werden.

Die Messung muss vorab in einer Testdurchführung exakt beschrieben sein. Insbesondere muss der Ablauf der Messung standardisiert und dementsprechend die Testgütekriterien für Messungen nachgewiesen sein. Danach werden die protokolliert Messungen an der LWS und der HWS, inklusive der demographischen/anthropometrischen Daten (Alter, Geschlecht, Körpergröße und -gewicht) in einer Datenbank gesammelt. Eine anthropometrische Gewichtung der Teilkörpermasse des Rumpfes [6] – Darstellung des maximalen isometrischen Drehmoments als *maximalen isometrischen Drehmoments pro Kilogramm Oberkörpergewicht* – ist notwendig zur Körpergewichts- und -größenunabhängigen Interpretation der nach alters- und geschlechtspezifisch unterteilten Untergruppen.

Tab. 10 Ergebnisse der asymptotischen Signifikanz (p-Wert) des Kruskal-Wallis-Tests bei den Frauen

Altersdekade	Extension	Flexion	Rotation rechts	Rotation links	Lateralflexion rechts	Lateralflexion links
<20 Jahre	0,015	0,420	0,986	0,744	0,774	0,437
20–29 Jahre	0,073	0,013	0,249	0,004	>0,001	0,001
30–39 Jahre	0,001	>0,001	0,005	0,001	>0,001	>0,001
40–49 Jahre	0,912	0,002	0,019	0,001	>0,001	>0,001
50–59 Jahre	0,680	0,914	0,266	0,010	0,001	0,013
>59 Jahre	0,013	0,120	0,037	0,046	0,312	0,618

Tab. 11 Ergebnisse der asymptotischen Signifikanz (p-Wert) des Kruskal-Wallis-Tests bei den Männern

Altersdekade	Extension	Flexion	Rotation rechts	Rotation links	Lateralflexion rechts	Lateralflexion links
<20 Jahre	0,433	0,021	0,003	0,005	0,001	0,021
20–29 Jahre	0,046	0,047	>0,001	>0,001	>0,001	>0,001
30–39 Jahre	>0,001	>0,001	0,005	>0,001	>0,001	>0,001
40–49 Jahre	0,043	0,008	>0,001	>0,001	>0,001	>0,001
50–59 Jahre	0,040	0,013	0,014	0,015	0,107	0,051
>59 Jahre	0,305	0,026	0,050	0,090	0,026	0,071

Referenzwerte darstellen, Referenzwerte bestimmen und stochastisches Rauschen reduzieren

Mit dem Ansatz der Gleichverteilung der repräsentativen Verteilungsmerkmale an der Gesamtbevölkerung, erhält man nach statistischem Ermessen eine hohe Sicherheit, Ausreißer, Redundanzen und ande-

re Effekte des statistischen Rauschens zu reduzieren.

Die Analyse bestimmt sich analog zu den Forderungen der IFCC in den folgenden Teilschritten:

- Histogramm untersuchen auf
 - Ausreißer – Wert prüfen
 - Mehrgipfligkeit – Population unterteilen
 - Schiefe – 3

Übersichten

Tab. 12 Beispiel der Entwicklung der Referenzwerte im Jahresverlauf, Rotation rechts (30–39 Jahre)

Jahr	Frauen (n)	MW ± SD (Nm)	95% CI (Nm)	Männer (n)	MW ± SD (Nm)	95% CI (Nm)
1996	-	-	-	24	3,32±0,78	2,99–3,65
1997	43	2,10±0,41	1,98–2,23	68	3,38±0,57	3,24–3,51
1998	67	2,06±0,56	1,93–2,20	71	3,07±0,80	2,89–3,26
1999	71	1,94±0,52	1,81–2,06	81	3,02±0,61	2,89–3,16
2000	111	2,00±0,58	1,89–2,11	136	2,98±0,73	2,86–3,11
2001	115	1,98±0,57	1,87–2,08	174	3,10±0,82	2,98–3,22
2002	143	2,10±0,52	2,01–2,19	197	3,08±0,86	2,96–3,20
2003	173	2,06±0,66	1,96–2,16	222	2,99±0,78	2,89–3,09
2004	134	1,91±0,69	1,79–2,02	181	3,02±0,79	2,90–3,13
2005	65	1,85±0,59	1,70–1,99	64	3,06±0,80	2,86–3,25
Gesamt	922	2,00±0,59	1,97–2,04	1218	3,06±0,78	3,02–3,11

MW Mittelwert, SD Standardabweichung, 95% CI 95% Konfidenzintervall.

Tab. 13 Beispiel der Entwicklung der Referenzwerte im Jahresverlauf, Rotation links (30–39 Jahre)

Jahr	Frauen (n)	MW ± SD (Nm)	95% CI (Nm)	Männer (n)	MW ± SD (Nm)	95% CI (Nm)
1996	-	-	-	24	3,60±0,73	3,30–3,91
1997	43	2,03±0,64	1,83–2,23	68	3,50±0,72	3,32–3,67
1998	67	2,31±0,64	2,16–2,47	71	3,24±0,74	3,06–3,42
1999	71	1,99±0,52	1,87–2,11	81	3,13±0,65	2,98–3,27
2000	111	2,03±0,57	1,92–2,14	136	3,05±0,76	2,92–3,18
2001	115	1,95±0,57	1,85–2,06	174	3,19±0,84	3,06–3,31
2002	143	2,10±0,56	2,01–2,19	197	3,07±0,79	2,96–3,18
2003	173	2,07±0,71	1,96–2,17	222	3,03±0,80	2,92–3,13
2004	133	1,93±0,67	1,81–2,04	181	3,04±0,79	2,93–3,16
2005	65	1,96±0,58	1,81–2,10	64	3,14±0,83	2,93–3,35
Gesamt	921	2,03±0,62	1,99–2,07	1218	3,12±0,79	3,08–3,17

MW Mittelwert, SD Standardabweichung, 95% CI 95% Konfidenzintervall.

Tab. 14 Beispiel der Entwicklung der Referenzwerte im Jahresverlauf, Rotation rechts (40–49 Jahre)

Jahr	Frauen (n)	MW ± SD (Nm)	95% CI (Nm)	Männer (n)	MW ± SD (Nm)	95% CI (Nm)
1996	-	-	-	10	3±0,4	2,71–3,29
1997	35	2,16±0,66	1,93–2,38	58	3,40±0,72	3,21–3,59
1998	67	2,20±0,75	2,02–2,38	90	3,15±0,68	3,01–3,30
1999	67	1,90±0,61	1,75–2,05	67	3,03±0,69	2,86–3,20
2000	85	2,08±0,71	1,92–2,23	129	3,02±0,64	2,91–3,13
2001	144	1,97±0,52	1,89–2,06	148	3,13±0,69	3,02–3,24
2002	145	1,95±0,58	1,85–2,04	189	3,03±0,73	2,93–3,14
2003	237	1,98±0,64	1,90–2,07	212	2,77±0,81	2,66–2,88
2004	141	1,86±0,58	1,77–1,96	171	3,00±0,71	2,89–3,11
2005	78	1,80±0,56	1,67–1,93	73	2,94±0,72	2,77–3,11
Gesamt	999	1,97±0,62	1,93–2,01	1147	3,01±0,73	2,97–3,05

MW Mittelwert, SD Standardabweichung, 95% CI 95% Konfidenzintervall.

- Konfidenzintervalle (insbesonder 95% Konfidenzintervall) bestimmen
- Referenztabelle erstellen mit Mittelwert, Standardabweichung und Fallzahl in jeder Stratifikationsgruppe

Beispiel einer als A-posterior-Stichprobe entwickelten Referenzdatenbank

Harter [7] hat die Entwicklung von Referenzdaten in einer A-posterior-Stichprobe für die Messung der Rumpfkraft über

einen Zeitraum von 10 Jahren betrachtet und deren Änderung vorgestellt. Diese sind in **Tab. 4, 5, 6, 7, 8, 9** mit den Ergebnissen der Varianzanalyse aufgeführt. Die Messungen erfolgten an der Geräteserie der FPZ Systems entwickelt von der Fa. Schnell GmbH/ Peutenhausen, einem in der Therapieszene üblichen Mess- und Analysesystem. Die Messungen erfolgten unter standardisierten Rahmenbedingungen. Dies wurde durch eine spezifische praktische und theoretische Ausbildung der Therapeuten bei der Deutschen Gesellschaft für Muskuloskeletale Medizin e. V. (DGMSM/FAC e. V.) sichergestellt.

Die dargestellten Verteilungen wurden konventionsgemäß mit folgenden Angaben beschrieben:

- MW: Mittelwert aller Messwerte in der jeweiligen alters- und geschlechtspezifischen Untergruppe
- SD: Standardabweichung
- 95% CI: 95% Konfidenzintervall (konventionelle Vorgabe der IFCC zur Darstellung des Mittelwerts)
- OWA: Ergebnis der Varianzanalyse über die Änderung der jährlichen Werte im Bemessungszeitraum
- Änd./Jahr: durchschnittliche Änderung des Mittelwerts pro Jahr
- KWT: Ergebnis des Kruski-Wallis-Tests zur Darstellung der Veränderung der Stichprobe im 10-Jahres-Zeitraum

Aus den Ergebnissen (**Tab. 4, 5, 6, 7, 8, 9**) ist unschwer zu erkennen, dass – um hierbei eine statistische belastbare Fallzahl zu erlangen – aufgrund der zufälligen Verteilung und der Definition der Auswahl erheblich höhere Fallzahlen realisiert werden müssen. Zudem zeigen sich in den tatsächlichen repräsentativen Verteilungsmerkmalen der bundesdeutschen Bevölkerung (**Tab. 2 und 3**) doch andere demografische Merkmale als die in einer solchen A-posterior-Stichprobe, wie sie durch das zufällige Sammeln der Daten im multizentrischen Datenerfassung gewonnen wurde.

Die Interpretation von repräsentativen Referenzdaten stellt ein nicht zu unterschätzendes Plausibilitätsproblem dar, welches über die Beurteilung der statistischen Randbedingungen hinausgeht. Der Idealfall ist die Entwicklung einer wie vorhergehend beschriebenen A-priori-Stich-

probe. Bei der Darstellung von Referenzdaten über eine – wie hier im Folgenden vorgestellten – A-posterior-Stichprobe wird die Auswahl zwar ebenso über vordefinierte Kriterien [15] festgelegt, bei der zugrundeliegenden Rekrutierung von Besuchern der einzelnen (Rücken-)Therapieeinrichtungen in der multizentrischen Erfassung als Referenzstichprobe lässt sich jedoch das Risiko einer stochastischen Verzerrung (motivationale Vorauswahl) nicht ausschließen. Auch bei vorliegender Normalverteilung der einzelnen Referenzstichproben aus □ Tab. 4, 5, 6, 7, 8, 9 (Ergebnisse des Kolmogorov-Smirnow-Tests beim Verfasser einzusehen) wurde zum Darstellen der Veränderung der einzelnen alters- und geschlechtsspezifischen Stichproben der gegen Verletzung der Randbedingungen wesentlich robustere nonparametrische Kruski-Wallis-Test gewählt (□ Tab. 10, 11). Er zeigt signifikant, dass die Referenzwerte im Jahresverlauf variieren.

Insbesondere die Veränderung der Referenzwerte im 10-Jahres-Zeitraum rechtfertigen auch die Notwendigkeit einer sorgfältigen Rekrutierung der Versuchspersonen und eines regelmäßigen Upgrades der Daten (einzelne geschlechtspezifische Altersklassen im Jahresvergleich: □ Tab. 12, 13, 14, 15, 16, 17). Die weitestgehend beschwerdefreien Versuchspersonen, die über das o. g. Körpertonkonzept nach Israel [13] der Referenzstichprobe zugewiesen wurden, kamen aus definierten präventiven Gründen in die Therapieeinrichtungen zur Analyse. Damit ist nicht auszuschließen, dass hier auch schon eine entsprechende physische und mentale Vorbelaistung bei den Versuchspersonen bestand. Dies führte u. U. zu den in den Tabellen beschriebenen Varianzen in den einzelnen Werten (Standardabweichungen) und einer entsprechenden Inhomogenität der Varianzen.

Sehr ausgeprägt zeigte sich die Veränderung der Werte in der Transversal- und Frontalebene bei den Männern. Die Sagitalebene war offensichtlich nicht so stark betroffen (□ Tab. 12, 13, 14). Diese deutlichen Veränderungen können mit lebensstilbedingten Veränderungen im einseitigen und/oder massiv eingeschränkten Bewegungsverhalten zusammenhängen. Aber auch der moderne technische Kom-

Tab. 15 Beispiel der Entwicklung der Referenzwerte im Jahresverlauf, Rotation links (40–49 Jahre)

Jahr	Frauen (n)	MW ± SD (Nm)	95% CI (Nm)	Männer (n)	MW ± SD (Nm)	95% CI (Nm)
1996	-	-	-	10	3,51±0,28	3,31–3,70
1997	35	2,14±0,69	1,90–2,37	58	3,52±0,70	3,33–3,70
1998	67	2,40±0,76	2,22–2,59	90	3,42±0,65	3,29–3,56
1999	67	1,95±0,63	1,80–2,10	67	3,13±0,75	2,94–3,31
2000	85	2,10±0,66	1,95–2,24	129	3,10±0,67	2,98–3,22
2001	144	2,00±0,54	1,91–2,09	148	3,15±0,68	3,04–3,26
2002	145	1,91±0,60	1,81–2,01	189	3,09±0,73	2,98–3,20
2003	236	1,99±0,63	1,91–2,07	212	2,85±0,82	2,74–2,96
2004	141	1,92±0,60	1,82–2,02	171	3,07±0,74	2,96–3,19
2005	78	1,89±0,55	1,77–2,02	73	3,04±0,72	2,87–3,20
Gesamt	998	2,00±0,63	1,96–2,04	1147	3,10±0,75	3,06–3,15

MW Mittelwert, SD Standardabweichung, 95% CI 95% Konfidenzintervall.

Tab. 16 Beispiel der Entwicklung der Referenzwerte im Jahresverlauf, Lateralflexion rechts (40–49 Jahre)

Jahr	Frauen (n)	MW ± SD (Nm)	95% CI (Nm)	Männer (n)	MW ± SD (Nm)	95% CI (Nm)
1996	-	-	-	10	3,58±0,77	3,02–4,13
1997	31	2,57±0,72	2,30–2,83	53	3,32±0,67	3,14–3,5
1998	67	2,57±0,66	2,41–2,73	88	3,27±0,66	3,13–3,41
1999	60	2,20±0,72	2,01–2,38	67	3,09±0,8	2,9–3,28
2000	78	2,36±0,75	2,19–2,53	124	3,07±0,69	2,95–3,2
2001	121	2,35±0,65	2,24–2,47	143	3,1±0,64	3–3,21
2002	128	2,15±0,67	2,03–2,26	174	2,96±0,74	2,85–3,07
2003	178	2,24±0,63	2,15–2,34	196	2,89±0,76	2,78–2,99
2004	123	2,13±0,66	2,01–2,25	156	2,99±0,78	2,87–3,12
2005	70	2,14±0,67	1,98–2,30	71	2,83±0,67	2,68–2,99
Gesamt	856	2,26±0,68	2,22–2,31	1082	3,03±0,73	2,99–3,07

MW Mittelwert, SD Standardabweichung, 95% CI 95% Konfidenzintervall.

Tab. 17 Beispiel der Entwicklung der Referenzwerte im Jahresverlauf, Lateralflexion links (40–49 Jahre)

Jahr	Frauen (n)	MW ± SD (Nm)	95% CI (Nm)	Männer (n)	MW ± SD (Nm)	95% CI (Nm)
1996	-	-	-	10	3,97±0,83	3,38–4,57
1997	31	2,41±0,79	2,12–2,70	53	3,34±0,78	3,12–3,55
1998	67	2,75±0,78	2,56–2,94	88	3,31±0,71	3,16–3,47
1999	60	2,42±0,76	2,23–2,62	67	3,31±0,77	3,12–3,50
2000	78	2,48±0,78	2,30–2,66	124	3,13±0,73	3,00–3,25
2001	121	2,43±0,56	2,33–2,53	143	3,14±0,69	3,03–3,25
2002	128	2,28±0,67	2,17–2,40	174	3,05±0,77	2,94–3,17
2003	178	2,33±0,65	2,23–2,43	196	2,91±0,77	2,80–3,02
2004	123	2,19±0,67	2,07–2,31	156	3,02±0,74	2,90–3,14
2005	70	2,21±0,72	2,04–2,39	71	2,93±0,74	2,76–3,11
Gesamt	856	2,36±0,70	2,32–2,41	1082	3,09±0,76	3,05–3,14

MW Mittelwert, SD Standardabweichung, 95% CI 95% Konfidenzintervall.

fort reduziert die natürlich mehrachsige Bewegung im Alltag.

Es ist aber auch nicht auszuschließen, dass – wie oben schon erwähnt – stochastische Phänomene eine größere Varianzhomogenität bedingen. Hier wäre im Weiteren die Entwicklung qualitativ

hochwertiger A-priori-Stichproben mit Sicherheit statistisch belastbarer.

Fazit für die Praxis

Validé Messungen in Verbindung mit einem wertenden Referenzdatenkon-

zept sind eine sinnvolle Unterstützung in der Differenzialdiagnostik der funktionellen Störungen bei Patienten mit chronischen Rückenbeschwerden. Aus der Analyse der wirbelsäulenstabilisierenden Muskulatur lassen sich – in Verbindung mit einer Differenzialdiagnostik – mögliche funktionelle und strukturelle muskuläre Störungen bei Patienten mit chronischen Rückenbeschwerden erfassen [8]. Gleichzeitig lässt sich der Erfolg angemessener aktiver Therapieverfahren damit valide dokumentieren. Allerdings müssen zur inhaltlichen Glaubwürdigkeit definierte wissenschaftliche Kriterien überprüft und nachgewiesen sein:

- Zertifizierung nach dem Medizinproduktegesetz, insbesondere eine ausführliche klinische Bewertung für das jeweilige Messsystem (technische Geeignetheit)
- Ausführliche Beschreibung und Dokumentation der Herkunft der Referenzdaten (Stichprobencharakteristika) zur Beurteilung der Repräsentativität
- Exakte Beschreibung der Daten und Sicherstellung der Messdurchführung im Sinne der Qualitätssicherung (Systematische Transparenz)
- Methodische Präzisierung der
 - Interpretation der Messergebnisse im Sinne des pathogenetischen Prinzips,
 - therapeutischen Schlüsse zur aktiven Behandlung,
 - qualitativen und quantitativen Bewertung des Behandlungserfolgs/-ergebnisses im physiologischen und psychologischen Sinn [8].

Die Erfüllung beschriebenen Kriterien können zur medizinisch-wissenschaftlichen, aber ökonomisch transparenten Bewertung entsprechender Analyse- und Therapiekonzepte herangezogen werden. So kann einerseits der behandelnde Arzt die Güte und Tauglichkeit solcher Messungen für seine Diagnose beurteilen, andererseits können die Kostenträger ein klare, von wissenschaftlichen Kriterien geleitete Abgrenzung nützlicher und tauglicher Methoden betreiben.

Korrespondenzadresse

Dr. W.H. Harter

Hirschweg 16, 50226 Frechen
info@wulfram-harter.de

Interessenkonflikt. Der korrespondierende Autor gibt an, dass kein Interessenkonflikt besteht.

Literatur

1. Addison R, Schultz A (1980) Trunk strengths in patients seeking hospitalization for chronic low-back disorders. *Spine* 5/6:539–544
2. Airaksinen O, Brox JI, Cedraschi C et al (2006) On behalf of the COST B13 Working Group on Guidelines for Chronic Low Back Pain. European Guidelines For The Management Of Chronic Non-Specific Low Back Pain. *Eur Spine J* 15(Suppl. 2):S192–S300
3. Alonso M, Finn EJ (2000) Physik, 3 Aufl. Oldenburger Wissenschaftsverlag GmbH, München Wien Oldenburg, S 18–24
4. Bandy WD, Lovelace-Chandler V, McKittrick P, Bandy B (1990) Adaptation of skeletal muscle to resistance training. *J Orthop Sports Phys Ther* 12/6:248–255
5. Dybkaer R, Solberg HE (1987) International Federation of Clinical Chemistry (IFCC), Scientific Committee, Clinical Section Expert Panel on Theory of Reference Values (EPTRV) and International Committee for Standardization in Haematology (ICSH) Standing Committee on Reference Values -Approved Recommendations (1987) on the Theory of Reference Values Part 6. Presentation of observed values related to reference values. *J Clin Chem Clin Biochem* 25:657–662
6. Harter WH (1999) Die Dynamik des Hebens in der Kritik als Gefährdungsbedingung am Arbeitsplatz. Objektivierung der Belastung durch Heben anhand der Bewegung in die Rumpfextension. Dissertation an der Sozialwissenschaftlichen Fakultät der Georg August Universität zu Göttingen
7. Harter WH (2006) Überprüfung eines referenz-datenstützten Funktionskonzepts für die wirbelsäulenstabilisierende Muskulatur in einer 10-jährigen Längsschnittstudie – Wird der deutsche Rücken immer schwächer? Deutscher Kongress für Orthopädie und Unfallchirurgie, Vortrag
8. Harter WH (2010) Systematische Medizinische Trainingstherapie. *Man Med* 48(5):353–359
9. Hollmann W, Hettinger T (1980) Sportmedizin – Arbeits- und Trainingsgrundlagen, 2. Aufl. Stuttgart
10. Holmes JA, Damaser MS, Lehmann SI (1994) Erector spinae activation and movement dynamics about the lumbar spine in lordotic and kyphotic squat-lifting. *Spine* 17/3:327–334
11. Hultmann G, Nordin M, Saraste H, Ohlsen H (1993) Body Composition, endurance, strength, cross-sectional area and density of mm erector spinae in men with and without low back pain. *J Spinal Disord* 6/2:114–123
12. Kommission „Human-Biomonitoring“ des Umweltbundesamtes (1996) Konzept der Referenz- und Human-Biomonitoring-Werte in der Umweltmedizin. *Bundesgesundheitsblatt* 6/96:221–223
13. Israel S, Freiwald J, Engelhardt M (1995) Bewegungsindizierte Adaptation und Körpermöglichkeiten. Gesund und leistungsfähig – Idealnorm ist das Ziel. *TW Sport + Medizin* 7(1):45–49
14. Janda V (1994) Manuelle Muskelfunktionsdiagnostik, Ullstein & Mosby, Berlin
15. Mayer TG (1985) Using physical measurements to assess low back pain. *J Musculoskeletal Med* 6:44–59
16. McNeill T, Warwick D, Anderson G, Schultz A (1980) Trunk strengths in attempted flexion, extension and lateral bending in health subjects and patients with low-back disorders. *Spine* 5/6:529–538
17. Parkola R, Kujala U, Rytökoski U (1992) Response of the trunk muscles to training assessed by magnetic resonance imaging and muscle strength. *Eur J Appl Physiol* 65:383–387
18. Petitclerc CC, Solberg HE (1987) International Federation of Clinical Chemistry(IFCC), Scientific Committee, Clinical Section Expert Panel on Theory of Reference Values (EPTRV) -Approved Recommendations (1987) on the Theory of Reference Values Part 2. Selection of individuals for the production of reference values. *J Clin Chem Clin Biochem* 25:639–644
19. Poulsen OM, Holst E, Christensen JM (1997) International Union of Pure and Applied Chemistry, Clinical Chemistry Division Commission on Toxicology -Calculation and Application of Coverage Intervals for Biological Reference Values (Technical Report): A supplement to the approved IFCC recommendations (1987) on the theory of reference values. *Pure Appl Chem* 69(7):1601–1611
20. Rutherford OM, Jones DA, Round JM (1990) Long-lasting unilateral muscle wasting and weakness following injury and immobilisation. *Scand J Rehabil Med* 22:33–37
21. Sale DG, McComas AJ, MacDougall JD, Upton AR (1982) Neuromuscular adaptation, in human rheumatic muscles follow strength training and immobilization. *J Appl Physiol* 53:419–424
22. Suzuki N, Endo S (1983) A quantitative study of trunk muscle strength and fatigability in the low-back-pain syndrome. *Spine* 8/1:69–74
23. Solberg HE (1987) International Federation of Clinical Chemistry (IFCC), Scientific Committee, Clinical Section Expert Panel on Theory of Reference Values (EPTRV) and International Committee for Standardization in Haematology (ICSH) Standing Committee on Reference Values -Approved Recommendations (1987) on the Theory of Reference Values Part 5. Statistical treatment of collected reference values. Determination of Reference Limits. *J Clin Chem Clin Biochem* 25:645–656
24. Solberg HE (1994) Establishment and use of reference values. In: Burts CA, Ashwood ER (Hrsg) Textbook of clinical chemistry, 2. Aufl. New York, S 454–484
25. Uhlig H (2005) Neue qualitätsgesicherte Diagnos-tikansätze bei chronischen Rückenschmerzen. In: Harter WH, Schifferdecker Hoch F, Brunner HE, Denner A (Hrsg) Menschen in Bewegung. Multidimensionale Neuausrichtung in der Behandlung chronischer Rückenschmerzen. Deutscher Ärzte, Köln, S 1–21
26. World Health Organization (1994) Assessing human health risks of chemicals. Derivation of guidance values for health-based exposure limits. ICPS -Environmental Health Criteria 170. WHO-Geneva