*Bias in retrospective assessment of perceived dental treatment effects
when using the Oral Health Impact Profile, Qual Life Res. 2018 Mar;27(3):775-782.*

Publikationspromotion
zur Erlangung des akademischen Grades

Dr. med. dent.

an der Medizinischen Fakultät
der Universität Leipzig

eingereicht von:

Antje Erler geb. Remmler

Geburtsdatum / Geburtsort:

22.07.1982 in Werdau

angefertigt an / in:

Universität Leipzig

Medizinische Fakultät

Poliklinik für Prothetik und Werkstoffkunde

Poliklinik für Kinderzahnheilkunde und Primärprophylaxe

Betreuer:

Prof. Dr. Christian Hirsch

PD Dr. Daniel R. Reißmann

Beschluss über die Verleihung des Doktorgrades vom: 20.08.2019

Inhaltsverzeichnis

1. Einführung in die Thematik.....	2
2. Formatierte Publikation.....	11
3. Zusammenfassung.....	19
4. Literaturverzeichnis.....	24
5. Anlagen	
5.1. Darstellung des eigenen Beitrags.....	27
5.2. Selbstständigkeitserklärung.....	28
5.3. Lebenslauf.....	29
5.4. Publikationen.....	30
5.5. Danksagung.....	31

Einführung in die Thematik

Die Ergebnisse klinischer Diagnostik und das Gesundheitsempfinden des Patienten korrelieren nicht zwangsläufig. Es gilt, die Perspektive hinsichtlich objektiver Diagnostik und der Wahrnehmung des Patienten anzupassen [1, 2]. Die Begriffe „Lebensqualität“ und „gesundheitsbezogene Lebensqualität“ stehen dabei im Mittelpunkt der untersuchten Studien und gewinnen in sämtlichen gesellschaftspolitischen Strukturen und der Gesundheitsforschung zunehmend an Bedeutung.

Die World-Health-Organization (WHO) definiert Gesundheit als „Zustand vollkommen körperlichen, geistigen und sozialen Wohlbefindens und nicht nur in Abwesenheit von Krankheit und Schwäche“. Die Gesundheit ist also ein wesentlicher Bestandteil der Lebensqualität und bezieht sich auf die subjektive Sichtweise der Patienten [3]. Eine Verbesserung klinischer Symptome führt somit nicht zwangsläufig zum verbesserten Empfinden des Patienten. Ein Patient mit technisch einwandfreiem Zahnersatz kann weiterhin über Beschwerden klagen, wogegen ein anderer mit seinem, aus klinischer Sicht insuffizienten Ersatz, zufrieden ist. In wissenschaftlicher Hinsicht ist Lebensqualität ein mehrschichtiges Konzept. Extrinsische (Umwelt, Gesellschaft, Arbeitsbedingungen u.a.) und intrinsische (Persönlichkeit, Verhalten u.a.) Faktoren sind wichtige Elemente der allgemeinen Lebensqualität, welche gesundheitsbezogene und mundgesundheitsbezogene Lebensqualität (MLQ) miteinbeziehen [4]. In der Gesundheitsforschung werden verschiedene Messinstrumente angewandt, um die MLQ greifbar zu machen. Mit Hilfe von Fragebögen zur Lebensqualität werden wissenschaftliche Hypothesen zu evaluiert. Das international am häufigsten angewandte Messinstrument der MLQ ist der Oral Health Impact Profile (OHIP- 49), ein standardisierter Fragebogen bestehend aus 49 Fragen unterteilt in vier Domänen (Funktionseinschränkungen des Kauystems, orofaziale Schmerzen, dentofaziale Ästhetik, psychosozialer Einfluss der Mundgesundheit) [3, 5]. Es existiert neben der englischen Originalversion auch eine deutschsprachige Variante (OHIP-G49) [3], was einen internationalen Vergleich vereinfacht [6].

Bei der Auswertung der Fragebögen gewinnt man einen ersten visuellen Überblick zum Niveau eingeschränkter MLQ, indem man die Antworten mit dem Eindruck, den

man bei ähnlichen Patienten während einer Therapie hatte, vergleicht. Präzise Ergebnisse erhält man, indem die Fragen mit ihrer Häufigkeitsangabe zu einem Gesamtwert aufsummiert werden. Zur Interpretation dieses Gesamtwertes werden Referenzwerte herangezogen. Diese Referenzwerte erlauben eine Bewertung eines individuellen Patienten mit nicht - behandlungssuchenden Vergleichspersonen aus der Durchschnittsbevölkerung. Solche „typischen“ Summenwerte sind beispielsweise 0 für Personen ohne abnehmbaren Zahnersatz, d.h. gewöhnlich geben diese Personen keine Probleme an, ein Gesamtwert von 4 für Personen mit abnehmbaren Teilprothesen und ein Gesamtwert von 6 für Patienten mit Totalprothesen [3].

Analytische Studien untersuchen dabei die Korrelation zwischen einer Intervention und einem, durch geeignete Messgrößen definiertem, Ergebnis [7]. Beim Studiendesign sind zwei mögliche Zeitverläufe zu unterscheiden:

Prospektiv angelegte Studien überprüfen eine im Vorhinein festgelegte Hypothese und beginnen danach mit der Datensammlung. Die Daten werden gemäß der Hypothese erhoben, um den Zusammenhang von Intervention und Ergebnis (Outcome) herzuleiten und statistisch zu untermauern [8]. Im Gegensatz dazu verwenden retrospektive Studien vorhandenes Datenmaterial, welches aus Patientenakten oder Fragebögen stammt. Die Rückblickende Untersuchung geht von einem Ereignis aus und untersucht danach die Einflussgrößen [9].

Typische prospektive Studien sind die sogenannten Kohortenstudien (Longitudinalstudien oder Follow-Up Studien). Ausgewählte Patienten werden einer Behandlungsstrategie zugeordnet, der Behandlungserfolg ist aber noch unbekannt [7]. Ein Vorteil dieser Vorgehensweise ist die Datenqualität, da das Material auf die Anforderungen der Studie zugeschnitten ist. Allerdings kann die Ermittlung der Daten sehr viel Zeit in Anspruch nehmen und erfordert oft, gerade bei geringer Inzidenz, eine große Anzahl an Teilnehmern. Des Weiteren gibt es auch monetäre Nachteile gegenüber retrospektiv durchgeführten Studien [8].

Ein Beispiel für retrospektive Studien sind Fall-Kontroll-Studien, bei denen erkrankte mit nichterkrankten Patienten verglichen werden. Da das Endereignis zu Beginn der Studie bereits eingetreten ist, erhält man relativ schnell die gewünschten Ergebnisse. Der zeitliche und finanzielle Aufwand ist also deutlich geringer. Doch die Datenqualität kann aufgrund der nicht korrigierbaren Dokumentationstechnik leiden. Bei der Verwendung von Fragebögen, bei denen man Patienten nach zurückliegenden

Ereignissen befragt, können Erinnerungslücken (Recall-Bias) die Ergebnisse der Studie verzerren [8].

Auch die Verbindung beider Studiendesigns ist möglich. Bei „retrospektiven Kohortenstudien“ ist die Untersuchungsrichtung prospektiv, die Datenerfassung aber retrospektiv. Von Interesse bei dieser Studie ist die Gegenüberstellung des jeweiligen Studiendesigns und die Auswertung der Ergebnisse, welche prospektive und retrospektive Charakteristiken aufweisen.

Die retrospektive Betrachtung betrifft die Überprüfung, ob ein sogenannter „Respond Shift“ auftritt. Response Shift ist die Reaktion auf einen veränderten Behandlungszustand und kann als „positives Phänomen“ des Therapieverlaufs gewertet werden. [10]. Damit soll der Beweis erbracht werden, dass signifikante Ergebnisse in Bezug auf die MLQ sowohl durch prospektive, als auch durch retrospektive Betrachtung innerhalb einer Studie möglich sind.

Mit Behandlungsbeginn wurden die Patienten über die Durchführung der Studie aufgeklärt und um Einverständnis durch eigenhändige Unterschrift gebeten. Fühlten sich die Patienten ausreichend aufgeklärt und bestätigten durch ihre Unterschrift die Teilnahme an dieser Studie, wurde ihnen der erste Baseline-Fragebogen (Pre - Test) ausgehändigt und erläutert.

Im Rahmen dieser Studien fand der OHIP-G unter Verwendung folgender Bausteine Anwendung: Funktionseinschränkungen des Kausystems, orofaziale Schmerzen, dentofaziale Ästhetik, psychosozialer Einfluss der Mundgesundheit [3]. Der Fragebogen soll einen schnellen Überblick zur psychosozialen Beeinträchtigung der Mundgesundheit geben. Die Fragen sollten vom Patienten selbstständig, in der Regel in weniger als 2-3 Minuten, beantwortet werden [4]. Die befragten Patienten wurden nach vier Wochen zur Nachkontrolle bestellt und aufgefordert einen Follow-Up - Fragebogen (Post - Test), sowie den Neubewertungsbogen (Then - Test) auszufüllen, um den Zustand zu Beginn der Therapie rückblickend einzuschätzen. Dabei wurden Post-Test und Then-Test zum gleichen Zeitpunkt ausgefüllt, Pre-Test und Then-Test beziehen sich auf den gleichen Zeitpunkt (vor Behandlungsbeginn). Die Fragen des OHIP-G49 beschreiben den Zeitraum des letzten Monats. Wie oft Probleme in Bezug auf die MLQ bestehen, wird in 5 Skalen („nie“=0, „kaum“=1, „ab und zu“=2, „oft“=3, „sehr oft“=4) angegeben [3, 4].

Ein weiteres Instrument, um die Veränderung der MLQ zu messen, ist die Verwendung der globalen Frage „Wie würden Sie Ihren Mundgesundheitszustand

einschätzen?“. Dabei stehen fünf Antwortkategorien zur Verfügung („sehr gut“, „gut“, „mittelmäßig“, „schlecht“). Eine weitere Frage mit den gleichen Antwortmöglichkeiten bezieht sich auf den allgemeinen Gesundheitszustand. Auch diese Methode kam jeweils zu Beginn der prothetischen Therapie und 4 bis 6 Wochen nach der Behandlung zum Einsatz.

Im Rahmen der Studie wurde bei 126 Patienten mit feststehendem Zahnersatz, partiellem Zahnersatz und Totalprothesen überprüft, ob während einer prothetischen Therapie unter Verwendung des standardisierten Fragebogens OHIP - 49 und einem gepaarten T-Test ein sogenannter Response Shift aufgetreten ist. Dabei wurde hinterfragt, ob eine Veränderung im Empfinden der MLQ im Verlauf der prothetischen Therapie festgestellt wird. Mit dieser Fragestellung zeigt sich, dass die Patienten ihre Werte und Ziele durch die Erfahrungen im Therapieverlauf ändern und neu definieren. Response shift ist also die Reaktion auf einen veränderten Behandlungszustand und kann als "positives Phänomen“ des Therapieverlaufs gewertet werden [10].

Die Patienten verändern nicht nur ihr Empfinden in Bezug auf die MLQ, sondern verändern offensichtlich auch ihre Sichtweise, was Lebensqualität für sie bedeutet. Diese Erkenntnis ermöglicht es, diesen positiven Effekt in den Behandlungsverlauf einfließen zu lassen, um nicht nur klinische Symptome zu behandeln, sondern auch das Wohlbefinden des Patienten zu verbessern.

Um bei der statistischen Auswertung den Response Shift sichtbar zu machen, wird die Differenz zwischen den OHIP-Summenwerten zu Beginn der Therapie (Pre-Test) mit denen der Neubewertung (Then-Test) gebildet [11]. Verbesserungen des Mundgesundheitszustandes nach prothetischer Therapie werden durch die Differenz zwischen Pre-Test und Post-Test angegeben. Um statistisch signifikante Ergebnisse bezüglich prothetischer Behandlungseffekte und Response Shift nachzuweisen, wurde der gepaarte t-Test angewendet.

Durch statistische Auswertung der Parameter und mit Hilfe des Then-Tests wird eine retrospektive Beurteilung des ersten Fragebogens (Pre - Test) mit dem letzten Fragebogen (Neubewertung) ermöglicht. Die einleitenden Worte unterscheiden sich natürlich vom ersten Fragebogen und sollen die Patienten anregen über ihre Mundgesundheit vor Behandlungsbeginn nachzudenken [12].

Die Methode geht davon aus, dass die Befragten beim Ausfüllen der Neubewertung ihren Anspruch an die prothetische Versorgung und die gewonnene Erfahrung nutzen, um sich, ausgehend vom Zeitpunkt vor der Behandlung, einzuschätzen [10].

Dabei sollten sie sich nicht erinnern, was im ersten Fragebogen angekreuzt wurde, sondern ihre damalige Empfindung, wie der Name schon sagt, neu bewerten.

Die Fragestellung bzw. Hypothese der Studie bezieht sich darauf, ob sich die Mittelwerte des Baseline - Bogens und die des Neubewertungsbogens unterscheiden. Diese Frage wurde mit dem oben aufgeführten gepaarten T-Test beantwortet, d. h., dass die Messwerte aus zwei Messungen an denselben Untersuchungseinheiten gewonnen wurden. Der T-Test überprüft dabei, ob ein Unterschied zwischen den Mittelwerten bedeutsam, also signifikant, ist. Wenn die Wahrscheinlichkeit, dass dieser Wert rein zufällig auftritt, kleiner als 5% ($p < 0,05$) ist, spricht man von Signifikanz, d. h. der vorliegende Mittelwertunterschied wäre bedeutsam. Mit dieser Studie sollte diese Hypothese bewiesen werden.

Als Beweis für den Respond Shift zeigt die rückblickende Neubewertung des Mundgesundheitszustandes eine schlechtere Selbsteinschätzung durch den Patienten (Then-Test) in Bezug auf den zu Beginn der Therapie vorhandenen Zahnersatz, als tatsächlich bei der initialen Befragung angegeben (Pre-Test).

Außerdem ist von Bedeutung, ob Behandlungseffekte prospektiv oder retrospektiv bestimmt werden sollen. Diese Fragestellung hat insbesondere Relevanz für die Translation von Forschungsergebnissen von zu erwartenden Behandlungsergebnissen in die Praxis und ist daher wichtig für klinische Entscheidungsfindungsprozesse. Während Effekte in klinischen Studien in der Regel prospektiv über den Vergleich einer Messung vor Intervention und einer Messung im Anschluss erfolgt, bewertet im klinischen Alltag der Patient den Erfolg der Behandlung nach deren Abschluss, in dem verglichen wird, ob sich eine bessere Lebensqualität und damit weniger Probleme, als vor der Behandlung einstellen [12]. Der Patient erinnert sich an den Zustand vor der Behandlung, was nur retrospektiv möglich ist. Mehrere Gründe sprechen dagegen, dass mit einer retrospektiven Methode erhobene Veränderungen vergleichbar mit prospektiv erhobenen Effekten sind. Betrachtet man den Zeitraum der Intervention, dann gibt es Hinweise, dass sich der interne Bezugsrahmen bzw. das Konzept von Lebensqualität des Patienten mit der Zeit verändert [10, 13]. Somit wären die Messungen von Baseline und Follow-Up nur eingeschränkt vergleichbar, da sich die Bewertungskriterien änderten. Dieser Effekt ist, wie oben beschrieben, als Response Shift bekannt und wird über mehrere Komponenten operationalisiert (z.B. Rekonzeptionalisierung, Repriorisierung und

Rekalibrierung) [14]. Bisherige Studien zeigten, dass solch ein Effekt bestehen könnte, welcher dazu führen würde, dass prospektiv erfasste Veränderungen gegenüber der retrospektiven Erfassung kleiner ausfallen. Dies könnte u.a. als eine Veränderung in den Bewältigungsstrategien (Coping) interpretiert werden [15-17]. Unter diesen Umständen wäre eine retrospektive Erfassung vorzuziehen. Diese weist aber auch erhebliche methodische Herausforderungen auf. Neben der offensichtlichen Fehlermöglichkeit bei der retrospektiven Erfassung, dass sich Patienten wahrscheinlich nicht mehr exakt an die Probleme und deren Häufigkeit vor der Behandlung erinnern können (Recall – Bias), kann die Einschätzung auch durch weitere Kriterien beeinflusst werden [18].

Einige Ergebnisse deuten darauf hin, dass der Effekt der Recall-Bias bei Anwendung der Then-Test-Methode vernachlässigbar und das Messmodell über die Zeit einigermaßen stabil ist [19]. Entsprechend können Respond-Shift-Effekte bei retrospektiver Bewertung nicht ausreichend durch Erinnerungseffekte der Patienten, oder Veränderungen der eigenen Bewertungsmaßstäbe und angewandter Bewältigungsstrategien, erklärt werden. Deshalb sollte man die psychologische Sichtweise der implizierten Veränderungstheorie und der Kognitiven Dissonanz in Betracht ziehen [20]. Die Patienten könnten den Zustand vor der Behandlung impliziert wiedergeben, basierend auf dem aktuellen Zustand und der empfundenen Veränderung hinsichtlich des Therapieergebnisses bzw. durch Inkongruenzen zwischen der erwarteten und der beobachteten Veränderung (Kognitive Dissonanz). Dieses Phänomen der implizierten Veränderungstheorie hätte aber zur Folge, dass die Einschätzung des Zustands vor der Behandlung nur eine Projektion ist und sich nicht auf die konkrete Situation zu diesem Zeitpunkt bezieht [21, 22]. Bisher gibt es bei zahnmedizinischen Patienten keine Ergebnisse, wie sich potentielle Effekte von Response Shift oder der implizierten Veränderung in Abhängigkeit der wahrgenommenen Veränderung auf die Erfassung der Lebensqualität auswirken. Es wäre zu erwarten, dass Patienten, welche eine Verbesserung wahrnehmen, nach der Behandlung auch eher besser mit den Problemen umgehen können (besseres Coping), aber mehr Probleme zum Ausgangszustand implizieren. Umgekehrt würden Patienten, welche eine Verschlechterung wahrnehmen, schlechter mit den Problemen zurechtkommen (schlechteres Coping), aber weniger Probleme zum Ausgangszustand implizieren [21, 23]. Beide Phänomene stimmen dahingehend überein, dass die retrospektive Beurteilung des Ausgangszustandes durch

Wahrnehmungen während der Behandlung rekonstruiert wird. Sie unterscheidet sich jedoch im jeweiligen Gesundheitszustand.

Die Theorie der kognitiven Dissonanz basiert auf der Annahme, dass die Patienten einen Zusammenhang zwischen ihren Erwartungen und der gegenwärtigen Situation suchen [23]. Die Patienten sind bestrebt diese Dissonanz zu reduzieren, indem sie ihre Einstellungen, Überzeugungen und Handlungen anpassen. Als Folge von fehlender Übereinstimmung zwischen der Erwartung und der wahrgenommenen Veränderung während der Behandlung, können Dissonanzen weiter bestehen. Um die Dissonanz zu beheben, können die Patienten die Ausgangssituation rückwirkend neu beurteilen, um die behandlungsinduzierte Veränderung in Übereinstimmung mit dem vermuteten Nutzen zu bringen. Demnach stellt die retrospektive Bewertung eine Verbindung zwischen der wahrgenommenen Veränderung und den erwarteten Verbesserungen in Bezug auf die Baseline-Bewertung her. Im Gegensatz dazu hängt die implizite Veränderungstheorie nicht notwendigerweise von einer Diskrepanz zwischen wahrgenommenem und erwartetem Nutzen ab. Die Patienten beurteilen ihren gegenwärtigen Zustand und die Dinge, die sich geändert haben könnten, um ein retrospektives Urteil zu fällen [21]. Darüber hinaus wurde die globale Bewertung der Veränderung der Mundgesundheit der Patienten im Vergleich zum Vor- und Nachbehandlungsstatus (Pre - Test, Post -Test) zur Nachkontrolle erfasst, mit der Frage „Wie hat sich Ihre Mundgesundheit seit Behandlungsbeginn verändert?“ (global transition question), mit den Antwortmöglichkeiten „viel verbessert“, „ein wenig verbessert“, „nicht verändert“, „ein wenig verschlechtert“, „viel verschlechtert“. Dadurch kann man die wahrgenommene Veränderung in Abhängigkeit von der Veränderung bei der retrospektiven Erhebung darstellen, also die Abhängigkeit des Respond Shift von der wahrgenommenen Veränderung.

Wir stellten also die Hypothese auf, dass der Respond Shift mit der globalen Bewertung der Veränderung der mundgesundheitsbezogenen Lebensqualität in Zusammenhang steht und konnten diese auch bestätigen.

Des Weiteren wurde untersucht, ob die retrospektive Baseline-Bewertung mehr mit der initialen Baseline-Bewertung oder dem Follow-Up zusammenhängt, was das Auftreten der implizierten Veränderungstheorie bzw. der Kognitiven Dissonanz erklären würde. Ziel dieser Studie war es daher, zu untersuchen, ob eine retrospektive Bewertung der mundgesundheitsbezogenen Lebensqualität bei

Verwendung des OHIP für die implizite Veränderungstheorie und Kognitive Dissonanz anfällig ist.

Referenzen

1. Armellini, D.B., et al., *Effect of removable partial dentures on oral health-related quality of life in subjects with shortened dental arches: a 2-center cross-sectional study*. Int J Prosthodont, 2008. **21**(6): p. 524-30.
2. de Oliveira, C.M. and A. Sheiham, *Orthodontic treatment and its impact on oral health-related quality of life in Brazilian adolescents*. J Orthod, 2004. **31**(1): p. 20-7; discussion 15.
3. John, M.T., D.L. Patrick, and G.D. Slade, *The German version of the Oral Health Impact Profile--translation and psychometric properties*. Eur J Oral Sci, 2002. **110**(6): p. 425-33.
4. John, M.T., et al., *Oral health-related quality of life in Germany*. Eur J Oral Sci, 2003. **111**(6): p. 483-91.
5. Slade, G.D. and A.J. Spencer, *Development and evaluation of the Oral Health Impact Profile*. Community Dent Health, 1994. **11**(1): p. 3-11.
6. Allison, P., et al., *A cross-cultural study of oral health values*. J Dent Res, 1999. **78**(2): p. 643-9.
7. Stengel, D., *Statistik und Aufbereitung klinischer Daten*. 2011: Thieme.
8. Weiß, C., *Basiswissen Medizinische Statistik*. 2013: Springer.
9. Lukas-Nülle, M.e.a., *Retrospektive versus prospektive Krankheitskostenstudien – Eine Analyse am Beispiel chronischer Schmerzpatienten*. Gesundh ökon Qual manag 2007: p. 247-252.
10. Schwartz, C.E., et al., *The clinical significance of adaptation to changing health: a meta-analysis of response shift*. Qual Life Res, 2006. **15**(9): p. 1533-50.
11. Broberger, E., M. Sprangers, and C. Tishelman, *Do internal standards of quality of life change in lung cancer patients?* Nurs Res, 2006. **55**(4): p. 274-82.
12. Visser, M.R., F.J. Oort, and M.A. Sprangers, *Methods to detect response shift in quality of life data: a convergent validity study*. Qual Life Res, 2005. **14**(3): p. 629-39.
13. Ring, L., et al., *Response shift masks the treatment impact on patient reported outcomes (PROs): the example of individual quality of life in edentulous patients*. Health Qual Life Outcomes, 2005. **3**: p. 55.
14. Wilson, I.B., *Clinical understanding and clinical implications of response shift*. Soc Sci Med, 1999. **48**(11): p. 1577-88.
15. Folkman, S. and R.S. Lazarus, *The relationship between coping and emotion: implications for theory and research*. Soc Sci Med, 1988. **26**(3): p. 309-17.
16. Hedlund, M., et al., *Coping strategies, health-related quality of life and psychiatric history in patients with aneurysmal subarachnoid haemorrhage*. Acta Neurochir (Wien), 2010. **152**(8): p. 1375-82; discussion 1382.
17. Yang, H.C., B.M. Brothers, and B.L. Andersen, *Stress and quality of life in breast cancer recurrence: moderation or mediation of coping?* Ann Behav Med, 2008. **35**(2): p. 188-97.

18. McPhail, S. and T. Haines, *Response shift, recall bias and their effect on measuring change in health-related quality of life amongst older hospital patients*. Health Qual Life Outcomes, 2010. **8**: p. 65.
19. Reissmann, D.R., et al., *Impact of response shift on the assessment of treatment effects using the Oral Health Impact Profile*. Eur J Oral Sci, 2012. **120**(6): p. 520-5.
20. Schwartz, C.E. and M.A. Sprangers, *Guidelines for improving the stringency of response shift research using the thentest*. Qual Life Res, 2010. **19**(4): p. 455-64.
21. Norman, G., *Hi! How are you? Response shift, implicit theories and differing epistemologies*. Qual Life Res, 2003. **12**(3): p. 239-49.
22. Festinger, L., *A theory of social comparison processes*. Hum Relat, 1954. **7**: p. 117-140.
23. Festinger L, C.J., *Cognitive consequences of forced compliance*. J Abnorm Psychol, 1959. **58**: p. 203-210.

Bias in retrospective assessment of perceived dental treatment effects when using the Oral Health Impact Profile

Daniel R. Reissmann¹ · Antje Erler² · Christian Hirsch³ · Ira Sierwald⁴ · Carolina Machuca⁵ · Oliver Schierz²

Accepted: 14 October 2017
© Springer International Publishing AG 2017

Abstract

Purpose Aim of this exploratory study was to investigate whether a retrospective assessment of oral health-related quality of life (OHRQoL) using the Oral Health Impact Profile (OHIP) is susceptible to bias such as implicit theory of change and cognitive dissonance.

Methods In this prospective clinical study, a sample of 126 adult patients (age 17–83 years, 49% women) requiring prosthodontic treatment was consecutively recruited. The OHRQoL was assessed using the 49-item OHIP at baseline and at follow-up. Additionally, patients were asked at follow-up to retrospectively rate their oral health status at baseline (retrospective pretest or then-test) and the change in oral health status using a global transition question. Furthermore, patients' ratings of overall oral health and general health were used as validity criteria for the OHRQoL assessments.

Response shift was calculated as the difference between the initial and retrospective baseline assessments.

Results Baseline and retrospective pretest did not differ substantially in terms of internal consistency and convergent validity. Response shift was more pronounced when patients perceived a large change in OHRQoL during treatment. Retrospective pretests were more highly correlated with the baseline than with the follow-up assessment.

Conclusion Findings suggest that retrospective assessments of OHRQoL using the OHIP-49 are susceptible to bias. Cognitive dissonance is more likely to appear as a source of bias than implicit theory of change.

Keywords OHRQoL · Treatment effect · Assessment · Response shift · OHIP

Daniel R. Reissmann and Antje Erler have contributed equally to this work and are considered as joint first authors.

✉ Daniel R. Reissmann
d.reissmann@uke.de

¹ Department of Prosthetic Dentistry, Center for Dental and Oral Medicine, University Medical Center Hamburg-Eppendorf, Martinistrasse 52, 20246 Hamburg, Germany

² Department of Prosthodontics and Materials Science, School of Dentistry, University of Leipzig, Leipzig, Germany

³ Department of Pediatric Dentistry, School of Dentistry, University of Leipzig, Leipzig, Germany

⁴ Department of Orthodontics, Dentofacial Orthopedics and Pedodontics, Center for Dental and Craniofacial Sciences, Charité – Universitätsmedizin Berlin, Berlin, Germany

⁵ Academic Unit of Dental Public Health, School of Clinical Dentistry, University of Sheffield, Sheffield, UK

Introduction

Patient-reported outcomes assessing health status are important measures in medicine and dentistry, with health-related and oral health-related quality of life (QoL) being commonly used constructs to assess patients' perceptions [1]. These measures allow not only to determine disease-related burdens but also to evaluate changes in QoL due to treatment effects and to compare different treatment options based on the patient perspective. Therefore, they are essential for both epidemiology and clinical decision-making.

To measure changes in clinical studies, QoL instruments are administered in at least two occasions, before and after the intervention, providing information whether or not the health status has improved. A key requirement is that the questionnaire measures the same construct on all occasions, but research has shown that this assumption might be violated, preventing from the simple translation

of the difference in the instrument's score between baseline and follow-up into the change during the intervention [2]. Patients might change their internal standards and ways of coping with oral and general health problems during the intervention [3–5], a phenomenon known as “response shift” [6]. It can be operationalized as reconceptualization, reprioritization, and recalibration of the measurement model [7]. To circumvent the response shift phenomenon, particularly changes in reference values [8] or when an initial baseline assessment is not possible [9], a common approach is a retrospective pretest–posttest design, or the then-test. The then-test retrospectively assesses the baseline status at the moment of the follow-up. Patients would rate their situation before the intervention applying the same internal standards and coping styles as for follow-up. Thus, response shift has been operationalized as the difference between the baseline and the then-test scores. Interestingly, this approach is well comparable to the determination of changes due to treatment in clinical practice, when patients are asked after the intervention to indicate whether and how much their situation has changed after the intervention when compared to the situation at baseline. Several studies in medicine and dentistry have shown considerable response shift effects when using this method [10–13], even though evidence regarding magnitude and direction of the effect is still inconclusive.

However, the then-test approach due to its retrospective nature is susceptible to several methodological challenges such as recall bias [14, 15]. It is not clear whether patients can accurately rate their situation in the past, especially for large time periods between baseline and follow-up assessments. Furthermore, the then-test might be prone to psychological phenomena that would act as confounders in the retrospective assessment, as cognitive dissonance and implicit theory of change [16–18] inducing changes in reported QoL without actually affecting internal standards. Both phenomena have in common that the retrospective judgment of the initial status is reconstructed affected by impressions during the intervention, but they differ in the health status based on the retrospective rating.

The theory of cognitive dissonance is based on the assumption that individuals seek consistency between their expectations and their reality [16]. This theory proposes that people have a motivational drive to reduce dissonance by adapting their attitudes, beliefs, and actions. Dissonances after interventions can occur as a result of inconsistencies between believed benefits and perceived changes. To resolve the dissonance, patients may retrospectively readjust the baseline assessment to bring the intervention-induced change into correspondence with the believed benefit. Accordingly, the retrospective baseline rating would be a function of the initial baseline status and the relationship between perceived improvements and expected benefits.

In contrast, the implicit theory of change does not necessarily depend on a discrepancy between perceived and expected benefits. Patients assess their present state and then ask how things might have changed to produce a retrospective judgement [17]. Thus, patients rate their status at follow-up and infer what their baseline status must have been based on their impressions during the intervention.

For oral health-related QoL (OHRQoL) when using the Oral Health Impact Profile (OHIP) [19], the most popular instrument in dentistry for OHRQoL assessment, some evidence suggests that the effect of recall bias when applying the then-test method is negligible [11], and the measurement model is reasonably stable over time, i.e., is longitudinally invariant [20]. Accordingly, response shift effects when assessed retrospectively cannot be sufficiently explained by memory effects or by changes in patients' internal standards and/or applied coping styles. Therefore, implicit theory of change and cognitive dissonance are proposed as alternatives [21].

We hypothesized that the presence and amount of response shift should be related to the overall perceived change in OHRQoL during treatment. Moreover, whether the retrospective baseline assessment is more related to the initial assessment or to the follow-up assessment could be interpreted as a possible explanation for either cognitive dissonance or implicit theory of change, respectively.

Thus, the aim of this exploratory study was to investigate whether a retrospective assessment of OHRQoL when using the OHIP is susceptible to implicit theory of change and cognitive dissonance.

Material and method

Study design, setting, and participants

In this prospective clinical study, a sample of 126 adult patients (age 17–83 years, 49% women) requiring prosthodontic treatment was consecutively recruited at the Department of Prosthodontics and Material Sciences at the University of Leipzig. All patients under 16 years of age or those who were not able (physically or mentally) to complete a questionnaire were excluded from the study.

The study has been reviewed by the Institutional Review Board (Reg. No.: 070-2005) of the Medical School, University of Leipzig, Leipzig, Germany. Informed consent was obtained from all individual participants included in the study. For more details regarding study design and patient recruitment see Reissmann et al. [11].

Assessment of OHRQoL and perceived oral and general health

The OHRQoL was assessed using the validated German 49-item version of the OHIP [19, 22] with a recall period of one month. Each OHIP item asks a patient to rate how often a specific impact on OHRQoL has been experienced within the last month, with possible responses on a five-point ordinal rating scale ("never" = 0, "hardly ever" = 1, "occasionally" = 2, "fairly often" = 3, "very often" = 4). OHRQoL is operationalized as the simple sum of all 49 OHIP items with potential summary scores ranging from 0 to 196 points. Higher scores indicate more problems, and hence lower OHRQoL.

Likewise, patients were asked to rate their overall oral health and their general health using a five-point ordinal rating scale ("excellent" = 4, "very good" = 3, "good" = 2, "moderate" = 1, "poor" = 0).

Patients completed the OHIP as a self-administered questionnaire before prosthodontic treatment (baseline) and 4–6 weeks after treatment (follow-up). During the follow-up assessment, patients were asked also to retrospectively rate their oral health status at baseline (then-test) using a specifically designed OHIP questionnaire. Response shift was calculated as the difference in OHIP summary scores between the baseline and the then-test.

Additionally, global rating of change in the patients' oral health comparing pre- and post-treatment status was assessed at follow-up by asking "How much has your oral health changed since before the treatment?" (global transition question) with possible responses on a five-point ordinal rating scale ("improved a lot," "improved a little," "about the same," "worsened a little," and "worsened a lot").

Statistical analyses

Our approach to investigate and compare prospective and retrospective assessment of changes in OHRQoL during prosthodontic treatment involved several steps.

First, we determined the number and proportion of patients for each response category of the global questions regarding perceived oral and general health for the three assessments (pre, post, and then-test). Statistical significance of pairwise comparisons was assessed using Wilcoxon signed-rank test. Additionally, measures for central tendency and variability (standard deviation; SD) of the OHIP summary score were computed for the three assessments, and differences of pairwise comparisons were tested for statistical significance by means of paired *t* tests.

Second, we assessed and compared reliability and validity of the then-test with the baseline and follow-up assessment. Therefore, we determined the instrument's internal consistency as a measure for reliability by computing

Cronbach's alpha [23] and average inter-item correlation. According to guidelines, alpha values of 0.70 or above indicate 'satisfactory' reliability [24]. For wide constructs such as OHRQoL, an average inter-item correlation of 0.15–0.20 is desirable [25]. Convergent validity was tested by computing Spearman's rank correlation coefficients for the correlation between the OHIP summary scores and the patients' global ratings of oral and general health for the different assessments. The magnitude of this coefficient was compared with guidelines suggesting that ≥ 0.3 represents a 'medium' effect and ≥ 0.5 represents a 'large' effect [26]. It was hypothesized that coefficients should be higher for the correlation between the OHIP summary scores and the patients' global ratings of oral health than with the global general health rating.

Third, we assessed whether differences between the baseline and the then-test (response shift) were related to the patients' perceptions of changes in oral health after the treatment. Therefore, we computed means and corresponding 95% confidence intervals (CI) for the differences in OHIP summary scores between the baseline and the then-test assessments stratified for global transition categories. Additionally, standardized effect sizes (Cohen's *d*) were computed for these categories. Guidelines suggest that an effect size of $d = 0.2$ is small, 0.5 is medium, and 0.8 is large [26]. This analysis was complemented by the computation of linear regression models with the difference in OHIP summary scores between baseline and then-test assessments (response shift) as the criterion variable and global transition categories as the predictor variable, assuming a linear relationship. Subsequently, the regression model was statistically adjusted for the potential confounders age and gender. We hypothesized that the amount of response shift is related to the change perceived by the patients when they utilize the perception of the treatment and/or implied improvements in the retrospective baseline rating.

Finally, we determined and compared the correlations between the OHIP summary scores of the baseline and then-test assessment for the categories of the global transition question. For the pairwise comparisons, Pearson product moment correlation coefficients were calculated. We hypothesized that correlations of the then-test with actual baseline and the follow-up assessment should differ, especially when perceived changes in oral health after the treatment are large. Higher coefficients for the correlation with the initial assessment might indicate that patients would remember the baseline status for the retrospective assessment, while higher coefficients for follow-up assessment would imply that patients utilized their oral health status at follow-up for the retrospective assessment. We interpreted stronger correlation of the then-test with the follow-up as indicator of implicit theory of change and stronger correlation of the then-test with the baseline as indicators of cognitive dissonance.

The amount of missing information was small with only 1.3% of all OHIP items. At the initial baseline assessment, 27 questionnaires (21.4%) had items with missing information. After treatment, 24 questionnaires (19.1%) for the follow-up assessment and 19 questionnaires (15.1%) for the retrospective baseline assessment had missing information. Missing answers in questionnaires with not more than 10% missing information were imputed using a regression method [27]. One questionnaire of the initial baseline assessment and two questionnaires of the follow-up assessment were excluded from further analysis as more than 10% of the items were unanswered.

All analyses were performed using the statistical software STATA/MP version 13.1 (StataCorp LP, College Station, TX, USA) with the probability of a type I error set at 0.05.

Results

Sample characteristics

Among the 126 patients in the study, approximately half were female (49.2%) and mean age was 54.8 (± 15.9) years. Before treatment, 60 patients (47.6%) had no or fixed dentures, further 60 patients (47.6%) had removable dental prostheses, and 6 patients (4.8%) had complete dentures.

At baseline, the majority of the patients rated their oral and their general health at least as "good" (61.7% and 74.8%, respectively; Table 1). After treatment, both oral and general perceived health improved, but only changes in oral health ratings were statistically significant. The retrospective

ratings of oral and general health were significantly worse than the follow-up ratings.

Differences between the baseline and the retrospective assessment were significant only for perceived oral health; on average the ratings were retrospectively worse. These findings for perceived oral health were in accordance with findings for OHRQoL. Highest OHIP summary score means were observed for the retrospective rating and lowest for the follow-up assessment, with statistically significant differences for all pairwise comparisons.

Psychometric properties of OHIP for different assessments

Psychometric properties of the OHIP assessed as reliability and validity did not differ substantially between the three

Table 2 Psychometric properties of OHRQoL assessment using the OHIP

	Reliability Cronbach's alpha	Validity	
		Perceived oral health Spearman's correlation coefficient	Perceived general health
Baseline	0.97	-0.57	-0.29
Follow-up	0.98	-0.42	-0.37
Retrospective	0.97	-0.66	-0.41

Correlations between OHIP-49 summary scores and the corresponding ratings for perceived oral and general health for each assessment

Table 1 Perceived oral and general health and OHRQoL for different assessments

	Assessment			Pairwise comparisons		
	Baseline <i>N</i> (%) or mean (SD)	Follow-up	Retrospective	BL-F	F-R	BL-R
<i>p</i> value						
Perceived oral health						
Excellent	0 (0.0)	4 (3.6)	3 (2.5)	<0.001	<0.001	0.013
Very good	9 (7.5)	17 (15.5)	9 (7.6)			
Good	65 (54.2)	75 (68.2)	45 (38.1)			
Moderate	31 (25.8)	13 (11.8)	44 (37.3)			
Poor	15 (12.5)	1 (0.9)	17 (14.4)			
Perceived general health						
Excellent	3 (2.5)	4 (3.5)	6 (5.1)	0.186	0.014	0.825
Very good	21 (17.7)	22 (19.5)	21 (17.8)			
Good	65 (54.6)	64 (56.6)	52 (44.1)			
Moderate	27 (22.7)	21 (18.6)	34 (28.8)			
Poor	3 (2.5)	2 (1.8)	5 (4.2)			
OHRQoL						
OHIP summary score	31.8 (28.5)	24.5 (24.5)	38.1 (32.0)	<0.001	<0.001	<0.001

assessments (Table 2). Internal consistency was considered 'satisfactory' according to guidelines [24] for all assessments, with values of Cronbach's $\alpha \geq 0.97$. OHIP summary scores were associated with global ratings of oral and general health in the expected direction, i.e., less oral health problems indicated as lower OHIP scores were associated with better global ratings. In general, correlation coefficients were higher for oral health than for general health ratings.

Results for the correlation between OHIP summary scores and patients' global ratings of oral health suggest that the coefficients represented 'large' effects for the initial and the retrospective assessment and only a 'medium' effect for the follow-up assessment. For patients' global ratings of general health, effects were 'medium' for the retrospective and follow-up assessment and 'small' for the baseline assessment. Although correlation coefficients were consistently higher for the retrospective assessment, differences were small and clinically not significant.

Perceived change and OHIP response shift

Most patients indicated that after treatment perceived oral health was better than before treatment (Table 3). Almost two-thirds (62.7%) rated their health status after treatment as "a lot better" and no one responded to the global transition question with "a lot worse."

The magnitude and direction of the differences in OHIP summary score between initial and retrospective baseline assessments (response shift) were related to patients' global rating (Table 3). Response shift was positive when patients indicated an improvement, i.e., OHIP summary scores of the retrospective assessments were higher than baseline, and negative when patients perceived deterioration or no change, i.e., baseline OHIP score was higher than the retrospective. Effect size indicates 'medium' effects for the categories "a lot better" and "a little worse," and 'small' effects for "a little better" and for "same." When a linear relationship between the global transition question and the magnitude of response

shift is assumed and computed in linear regression models, an increase of one category in the global rating is associated with a change in response shift of on average 7.4 OHIP points ($p < .001$). Results slightly changed (7.2 OHIP points, $p < .001$) after statistically controlling for the potential confounders age and gender.

Pairwise correlations of the OHIP summary scores between the three assessments revealed that the baseline assessment was substantially correlated to the follow-up with $r = .78$ in the whole sample. Furthermore, the then-test showed stronger correlation with the baseline ($r = .82$) than with the follow-up assessment ($r = .66$). When distinguishing between the global transition categories, the coefficient for the correlation between baseline and follow-up assessments was lowest for patients indicating "a lot better" for the global transition question, representing the largest improvement in perceived oral health during treatment (Table 3). When comparing the retrospective assessment with the baseline and follow-up, correlation coefficients differed only slightly for the categories "a little better" and "same" and were substantially stronger for the correlation with the baseline for the category "a lot better." This indicates that the retrospective assessment was more related to the baseline than to the follow-up and according to the proposed operationalization, this might be interpreted as cognitive dissonance.

Discussion

To the authors' knowledge, this is the first study in dentistry investigating whether the amount and direction of response shift is associated with the global perceived transition, and whether the retrospective pretest depends either on the initial baseline status or the perceptions at follow-up. Findings of the study suggest that while the initial baseline assessment and the then-test did not differ substantially in terms of internal consistency and convergent validity, response shift was more pronounced when patients perceived a large change

Table 3 OHIP response shift and inter-assessment correlations stratified by global change ratings

Global transition category	Subjects <i>N</i> (%)	Response shift		Pairwise correlations*		
		OHIP diff		BL-F	F-R	BL-R
		Mean (CI)	Effect size	Pearson correlation coefficient		
"a lot better"	72 (63.7)	9.4 (5.4; 13.3)	0.59	0.70	0.60	0.84
"a little better"	19 (16.8)	7.3 (-1.8; 16.5)	0.39	0.92	0.88	0.81
"same"	18 (15.9)	-4.8 (-11.1; 1.5)	-0.38	0.84	0.87	0.92
"a little worse"	4 (3.5)	-18.8 (-67.1; 29.6)	-0.62	-	-	-
"a lot worse"	-	-	-	-	-	-

*Sample size of $n=4$ for global transition category "a little worse" was deemed insufficient to compute correlations

BL baseline, F follow-up, R retrospective

after the treatment and the then-test was more strongly correlated with the baseline assessment than with the follow-up assessment.

Values for internal consistency and correlation with external standards of the baseline assessment and the then-test indicate that both assessments were reliable and valid. However, this does not necessarily mean that both assessed an identical construct. For the then-test, we used the retrospective global oral and general health ratings and these might also be prone to response shift effects. Thus, when applying guidelines for an instrument's reliability and validity, a retrospective OHRQoL assessment is reliable and valid but it might measure the baseline status from a different perspective than the initial assessment. One potential explanation might be that internal standards and applied coping styles had changed during the intervention.

Several studies have demonstrated that coping styles can have a direct or moderating effect on QoL [4, 5, 28–30], and intervention-induced changes in applied coping mechanisms are associated with increase in QoL [3]. Such changes in coping styles seem to be mainly due to recalibration [31], i.e., a change in patients' internal standards. Interestingly, while some consider recalibration as a form of a placebo effect [18], others interpret changes in coping styles as adaptation to illness, and hence as an intended and desirable treatment effect [15, 32, 33]. However, this might be true for chronically ill patients or for patients with newly diagnosed diseases such as cancer or injuries reporting substantially impaired QoL. This assumption is supported by studies showing that retrospectively assessed QoL is substantially better than initial QoL when the health status was deteriorating [12, 34], which could be interpreted as more adaptive and less maladaptive coping styles at follow-up.

Nonetheless, recalibration is more likely to occur in patients who report either improvement or deterioration on their condition than those who remain stable over time and this would result in same magnitude of recalibration but with different directions. For interventions with improved health status associated as in our study, retrospective assessments usually indicate worse QoL than assessed initially [10, 12]. But this would correspond to poorer coping at follow-up. We see no reason why patients should change their coping to less adaptive and/or more maladaptive style during the intervention. Therefore, it might be assumed that a change in applied coping styles cannot adequately explain the observed response shift phenomenon. Furthermore, a recent study has shown that the measurement model of the OHIP is longitudinally invariant, i.e., no substantial recalibration was observed [20].

In contrast, our findings indicate that the then-test is substantially related to the global change in perceived oral health during treatment. The better the patients perceived their oral health after treatment, the worse was OHRQoL

in the then-test, corresponding to a larger implied treatment effect. Accordingly, this suggests that patients rated their status at baseline retrospectively based on their experiences during and after treatment. The stronger correlation of the then-test with the baseline than with follow-up assessment also might indicate that patients remembered their baseline status and adopted their retrospective ratings based on the implied change during treatment. One potential explanation could be cognitive dissonance [16]. Cognitive dissonance has been related to effort justification and positive changes in health. Patients have invested efforts during the treatment and the results might be lower than expected for such an effort. Retrospectively assessing themselves as worse, they would justify this effort and compensate this dissonance. Since the retrospective pretest is based on the initial baseline status and not on the status at follow-up, no indicator for implicit theory of change was observed.

The comparability with the literature is limited as we are not aware of any study investigating the association of response shift with global perceived transition. However, one study has tested how the then-test is correlated with the baseline and the follow-up status of sleep function and HRQoL [35]. Findings of this study were in accordance to ours with higher correlation between the then-test and the baseline assessments than the follow-up assessment. In a systematic review on response shift effects, the authors observed substantial differences in the magnitude and the direction reported in the included studies [12]. This finding is comparable to our study as we found different effect sizes for response shift based on patients' global transition question. This relationship might explain the inconclusive findings in the systematic review. Some other reviews have also mentioned the effect of cognitive dissonance on ratings of satisfaction [36] and on treatment evaluation [37], but their results are not comparable to ours.

A strength of our study is the use of the OHIP [19], the most commonly applied and largely methodically investigated instrument to assess OHRQoL, available in several language versions, and fulfilling most of the criteria for health status and quality of life instruments as proposed by the Medical Outcomes Trust [20, 38–46]. Furthermore, sample size was deemed sufficiently large to detect clinically meaningful effects, and our patients can be considered as typical dental patients [47]. However, the dental interventions performed in our study were obviously highly effective in reducing perceived oral health problems, reflected in the global change ratings with almost two-thirds of the patients perceiving the situation after treatment a lot better. Although this is desirable from a clinical perspective, the possibility to investigate response shift in patients who perceived adverse treatment effects was limited. Despite some evidence suggests that the direction of response shift is related to whether patients

perceived either positive or adverse effects due to the intervention, the interpretation regarding adverse effects should be viewed with great caution. Since no control group was included and the allocation to the treatment group was obviously not randomized, the difference between the baseline and follow-up scores just represents a change in perceived oral health but cannot be directly translated into the treatment effect. Additionally, since this is an exploratory study, no further measures were included to detect cognitive dissonance or implicit theory of change. Further research on this topic is recommended.

We used global ratings of oral and general health as external criteria in the validity analyses. Even though such single questions about self-reported oral and general health that use a global rating have been well studied and widely used in clinical and population settings [48, 49], they are not equivalent to summary scores of comprehensive multi-item instruments for OHRQoL and HRQoL. When responding to a single question, patients can consider all aspects of oral and general health, can ignore aspects that are not relevant to them, and can differentially weigh the impact of the relevant aspects according to their individual perception. In contrast, for the simple sum of all items of a comprehensive instrument such as the OHIP, no weights are applied. Therefore, correlations between global ratings and OHIP summary score should not be very high but still substantial, as observed in our study.

Our findings have important clinical implications. The results of our study suggest that changes in OHRQoL retrospectively rated by patients after therapy are larger than when assessed prospectively. While some suggest to include both a prospective and a retrospective assessment to allow for comparisons and as a sensitivity analysis [15], our findings indicate that the retrospective assessment is biased by the patients' perceptions. Accordingly, results from prospective and retrospective assessments of treatment effects are not comparable.

Considering the limitations of our study, we suggest that cognitive dissonance is more likely a source of bias in retrospective OHRQoL assessments when using the OHIP than implicit theory of change. However, further studies are required to prove this hypothesis.

Acknowledgements The authors are grateful to Ms. Annett Schrock (University of Leipzig) for her help with data management.

Compliance with ethical standards

Conflict of interest The authors declare that they have no conflict of interest.

Informed consent Informed consent was obtained from all individual participants included in the study.

Research involving human participants All procedures performed in studies involving human participants were in accordance with the ethical standards of the institutional and/or national research committee and with the 1964 Helsinki Declaration and its later amendments or comparable ethical standards.

References

- Reissmann, D. R. (2016). Alignment of oral health-related with health-related quality of life assessment. *Journal of Prosthodontic Research*, 60(2), 69–71.
- Allison, P. J., Locker, D., & Feine, J. S. (1997). Quality of life: A dynamic construct. *Social Science & Medicine*, 45(2), 221–230.
- Kanitz, J. L., Pretzer, K., Reif, M., Voss, A., Brand, R., Warschburger, P., et al. (2011). The impact of eurythmy therapy on stress coping strategies and health-related quality of life in healthy, moderately stressed adults. *Complementary Therapies in Medicine*, 19(5), 247–255.
- Heydecke, G., Tedesco, L. A., Kowalski, C., & Inglehart, M. R. (2004). Complete dentures and oral health-related quality of life: Do coping styles matter? *Community Dentistry and Oral Epidemiology*, 32(4), 297–306.
- Kenne Sørensen, E., Ohlen, J., Jonsson, T., & Gaston-Johansson, F. (2007). Coping with recurrent breast cancer: Predictors of distressing symptoms and health-related quality of life. *Journal of Pain and Symptom Management*, 34(1), 24–39.
- Sprangers, M. A., & Schwartz, C. E. (1999). Integrating response shift into health-related quality of life research: A theoretical model. *Social Science & Medicine*, 48(11), 1507–1515.
- Schwartz, C. E., & Sprangers, M. A. (1999). Methodological approaches for assessing response shift in longitudinal health-related quality-of-life research. *Social Science & Medicine*, 45(11), 1531–1548.
- Nieuwkerk, P. T., Tollenaar, M. S., Oort, F. J., & Sprangers, M. A. (2007). Are retrospective measures of change in quality of life more valid than prospective measures? *Medical Care*, 45(3), 199–205.
- Watson, W. L., Ozanne-Smith, J., & Richardson, J. (2007). Retrospective baseline measurement of self-reported health status and health-related quality of life versus population norms in the evaluation of post-injury losses. *Injury Prevention*, 13(1), 45–50.
- Ring, L., Hofer, S., Houston, F., Harris, D., & O'Boyle, C. A. (2005). Response shift masks the treatment impact on patient reported outcomes (PROs): The example of individual quality of life in edentulous patients. *Health and Quality of Life Outcomes*, 3, 55.
- Reissmann, D. R., Remmler, A., John, M. T., Schierz, O., & Hirsch, C. (2012). Impact of response shift on the assessment of treatment effects using the Oral Health Impact Profile. *European Journal of Oral Sciences*, 120(6), 520–525.
- Schwartz, C. E., Bode, R., Repucci, N., Becker, J., Sprangers, M. A., & Fayers, P. M. (2006). The clinical significance of adaptation to changing health: A meta-analysis of response shift. *Quality of Life Research*, 15(9), 1533–1550.
- Ahmed, S., Mayo, N. E., Wood-Dauphinee, S., Hanley, J. A., & Cohen, S. R. (2005). The structural equation modeling technique did not show a response shift, contrary to the results of the then test and the individualized approaches. *Journal of Clinical Epidemiology*, 58(11), 1125–1133.
- McPhail, S., & Haines, T. (2010). Response shift, recall bias and their effect on measuring change in health-related quality of life amongst older hospital patients. *Health and Quality of Life Outcomes*, 8, 65.

15. Blome, C., & Augustin, M. (2015). Measuring change in quality of life: Bias in prospective and retrospective evaluation. *Value in Health*, 18(1), 110–115.
16. Festinger, L., & Carlsmith, J. M. (1959). Cognitive consequences of forced compliance. *Journal of Abnormal Psychology*, 58(2), 203–210.
17. Norman, G. (2003). Hi! How are you? Response shift, implicit theories and differing epistemologies. *Quality of Life Research*, 12(3), 239–249.
18. Wilson, I. B. (1999). Clinical understanding and clinical implications of response shift. *Social Science & Medicine*, 48(11), 1577–1588.
19. Slade, G. D., & Spencer, A. J. (1994). Development and evaluation of the Oral Health Impact Profile. *Community Dental Health*, 11(1), 3–11.
20. Reissmann, D. R., John, M. T., Feuerstahler, L., Baba, K., Szabo, G., Celebic, A., et al. (2016). Longitudinal measurement invariance in prospective oral health-related quality of life assessment. *Health and Quality of Life Outcomes*, 14(1), 88.
21. Schwartz, C. E., & Sprangers, M. A. G. (2010). Guidelines for improving the stringency of response shift research using the then-test. *Quality of Life Research*, 19(4), 455–464.
22. John, M. T., Patrick, D. L., & Slade, G. D. (2002). The German version of the Oral Health Impact Profile—translation and psychometric properties. *European Journal of Oral Sciences*, 110(6), 425–433.
23. Cronbach, L. J. (1951). Coefficient alpha and the internal structure of tests. *Psychometrika*, 16, 297–334.
24. Bland, J. M., & Altman, D. G. (1997). Cronbach's alpha. *British Medical Journal*, 314(7080), 572.
25. Clark, L. A., & Watson, D. (1995). Constructing validity: Basic issues in objective scale development. *Psychological Assessment*, 7(3), 309–319.
26. Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd edn). Hillsdale, NJ: Lawrence Erlbaum Associates.
27. John, M. T., LeResche, L., Koepsell, T. D., Hujoel, P., Migliorini, D. L., & Micheelis, W. (2003). Oral health-related quality of life in Germany. *European Journal of Oral Sciences*, 111(6), 483–491.
28. Pucheu, S., Conso, S. M., D'Auzac, C., Francois, P., & Issad, B. (2004). Do health causal attributions and coping strategies act as moderators of quality of life in peritoneal dialysis patients? *Journal of Psychosomatic Research*, 56(3), 317–322.
29. Taylor, J. L., Smith, P. J., Babyak, M. A., Barbour, K. A., Hoffman, B. M., Sebring, D. L., et al. (2008). Coping and quality of life in patients awaiting lung transplantation. *Journal of Psychosomatic Research*, 65(1), 71–79.
30. Yang, H. C., Brothers, B. M., & Andersen, B. L. (2008). Stress and quality of life in breast cancer recurrence: Moderation or mediation of coping? *Annals of Behavioral Medicine*, 35(2), 188–197.
31. Dempster, M., Carney, R., & McClements, R. (2010). Response shift in the assessment of quality of life among people attending cardiac rehabilitation. *British Journal of Health Psychology*, 15(Pt 2), 307–319.
32. Ubel, P. A., Peeters, Y., & Smith, D. (2010). Abandoning the language of "response shift": A plea for conceptual clarity in distinguishing scale recalibration from true changes in quality of life. *Quality of Life Research*, 19(4), 465–471.
33. Ubel, P. A., & Smith, D. M. (2010). Why should changing the bathwater have to harm the baby? *Quality of Life Research*, 19(4), 481–482.
34. Korfage, I. J., de Koning, H. J., & Essink-Bot, M. L. (2007). Response shift due to diagnosis and primary treatment of localized prostate cancer: a then-test and a vignette study. *Quality of Life Research*, 16(10), 1627–1634.
35. Meyer, T., Richter, S., & Raspe, H. (2013). Agreement between pre-post measures of change and transition ratings as well as then-tests. *BMC medical research methodology*, 13, 52.
36. Carr-Hill, R. A. (1992). The measurement of patient satisfaction. *Journal of Public Health Medicine*, 14(3), 236–249.
37. Homer, J. J., Sheard, C. E., & Jones, N. S. (2000). Cognitive dissonance, the placebo effect and the evaluation of surgical results. *Clinical Otolaryngology and Allied Sciences*, 25(3), 195–199.
38. Scientific Advisory Committee of the Medical Outcomes Trust. (2002). Assessing health status and quality-of-life instruments: attributes and review criteria. *Quality of Life Research*, 11(3), 193–205.
39. John, M. T., Reissmann, D. R., Schierz, O., & Allen, F. (2008). No significant retest effects in oral health-related quality of life assessment using the Oral Health Impact Profile. *Acta Odontologica Scandinavica*, 66(3), 135–138.
40. John, M. T., Reissmann, D. R., Szentpetery, A., & Steele, J. (2009). An approach to define clinical significance in prosthodontics. *Journal of Prosthodontics*, 18(5), 455–460.
41. Reissmann, D. R., John, M. T., & Schierz, O. (2011). Influence of administration method on oral health-related quality of life assessment using the Oral Health Impact Profile. *European Journal of Oral Sciences*, 119(1), 73–78.
42. Reissmann, D. R., Sierwald, I., Heydecke, G., & John, M. T. (2013). Interpreting one oral health impact profile point. *Health and Quality of Life Outcomes*, 11(1), 12.
43. John, M. T., Feuerstahler, L., Waller, N., Baba, K., Larsson, P., Celebic, A., et al. (2014). Confirmatory factor analysis of the Oral Health Impact Profile. *Journal of Oral Rehabilitation*, 41(9), 644–652.
44. Kieffer, J. M., & Hoogstraten, J. (2008). Item-order effects in the Oral Health Impact Profile (OHIP). *European Journal of Oral Sciences*, 116(3), 245–249.
45. Kieffer, J. M., Verrips, G. H., & Hoogstraten, J. (2011). Instrument-order effects: Using the Oral Health Impact Profile 49 and the Short Form 12. *European Journal of Oral Sciences*, 119(1), 69–72.
46. Sierwald, I., John, M. T., Durham, J., Mirzakhani, C., & Reissmann, D. R. (2011). Validation of the response format of the Oral Health Impact Profile. *European Journal of Oral Sciences*, 119(6), 489–496.
47. John, M. T., Reissmann, D. R., Feuerstahler, L., Waller, N., Baba, K., Larsson, P., et al. (2014). Factor analyses of the Oral Health Impact Profile: Overview and studied population. *Journal of Prosthodontic Research*, 58(1), 26–34.
48. Gill, T. M., & Feinstein, A. R. (1994). A critical appraisal of the quality of quality-of-life measurements. *Journal of the American Medical Association*, 272(8), 619–626.
49. Subramanian, U., Eckert, G., Yeung, A., & Tierney, W. M. (2007). A single health status question had important prognostic value among outpatients with chronic heart failure. *Journal of Clinical Epidemiology*, 60(8), 803–811.

Zusammenfassung der Arbeit

Diese Studie untersucht, ob eine retrospektive Beurteilung der Lebensqualität im Zusammenhang mit dem Mundgesundheitszustand (MLQ/ OHRQoL), unter Verwendung des Oral Health Impact Profile (OHIP), anfällig für Erinnerungsverzerrungen/ Recall-Bias ist, im Sinne der implizierten Veränderungstheorie und/ oder der kognitiven Dissonanz.

In wissenschaftlicher Hinsicht ist Lebensqualität ein mehrschichtiges Konzept. Extrinsische (Umwelt, Gesellschaft, Arbeitsbedingungen u.a.) und intrinsische (Persönlichkeit, Verhalten u.a.) Faktoren sind wichtige Bestandteile der allgemeinen Lebensqualität, welche gesundheitsbezogene und mundgesundheitsbezogene Lebensqualität (MLQ/ OHRQoL) einbeziehen [1]. Die Gesundheitsforschung nutzt verschiedene Messinstrumente, um die MLQ greifbar zu machen. Fragebögen zur Lebensqualität sind dabei ein wichtiges Instrument, um wissenschaftliche Hypothesen zu evaluieren. Das international am häufigsten angewandte Messinstrument der MLQ ist der Oral Health Impact Profile (OHIP- 49), ein standardisierter Fragebogen bestehend aus 49 Fragen [2, 3]. Es existiert neben der englischen Originalversion auch eine deutschsprachige Variante (OHIP-G49) [2], was einen internationalen Vergleich vereinfacht [4].

Die gesundheitsbezogene und mundgesundheitsbezogene Lebensqualität (MLQ/ OHRQoL) wird häufig genutzt, um die Patientenwahrnehmung darzustellen und darüber hinaus die Bewertung von Veränderungen der Lebensqualität auf Grund von Behandlungseffekten zu ermitteln [5]. Diese Erkenntnisse sind für die klinische Entscheidungsfindung von Bedeutung.

Um Veränderungen zu zeigen, wird der Fragebogen jeweils vor und nach der Therapie eingesetzt und gibt Aufschluss darüber, ob sich der (Mund-) Gesundheitszustand verbessert hat oder nicht.

Bei der Verwendung von diesen Fragebögen, als Instrument zur Messung der Lebensqualität, wird davon ausgegangen, dass der Bezugspunkt unverändert bleibt, d. h., dass die Einstellung eines Patienten gegenüber einem bestimmten Konstrukt stabil bleibt. Andernfalls können Änderungen als Reaktion auf bestimmte Variablen nicht interpretiert werden. Die Einstellungen sind jedoch nicht konstant: Sie variieren mit der Zeit, durch Erfahrungen und werden durch psychologische Phänomene wie

Anpassung, Bewältigungsstrategien und Erwartungen verändert [6]. Patienten können während einer Therapie ihre internen Standards zur Bewältigung oraler und allgemeiner Gesundheitsprobleme ändern, ein Phänomen, dass als Respond Shift bezeichnet wird [7]. Ein gängiger Ansatz, um das Respond-Shift-Phänomen zu umgehen, ist die Verwendung des Then-Test [8]. Der Then-Test beurteilt retrospektiv den Ausgangsstatus zum Zeitpunkt der Nachuntersuchung. Die Patienten bewerten rückblickend ihre Situation vor dem Eingriff unter Anwendung der selben internen Standards und Bewältigungsstile, wie für das Follow-Up. Daher wird Respond Shift als Differenz zwischen den Ausgangswerten und den dann erzielten Werten im Then-Test standardisiert. Der Then-Test ist jedoch auf Grund seines rückblickenden Charakters anfällig für verschiedene methodologische Herausforderungen, wie etwa Recall-Bias [9]. Darüber hinaus könnte der Then-Test für psychologische Phänomene anfällig sein, die bei der retrospektiven Beurteilung als Störfaktoren wirken, wie die kognitive Dissonanz und die implizite Veränderungstheorie. Diese können Änderungen in der zuvor berichteten Lebensqualität auslösen, ohne dass dies interne Standards beeinflusst [10-12]. Die Theorie der kognitiven Dissonanz basiert auf der Annahme, dass Individuen eine Übereinstimmung zwischen ihren Erwartungen und der Realität suchen. Sie sind bestrebt, bestehende Dissonanzen zu reduzieren, indem sie ihre Handlungen und Einstellungen anpassen [10]. Im Gegensatz dazu hängt die implizite Veränderungstheorie nicht unbedingt von einer Diskrepanz zwischen wahrgenommenem und erwartetem Nutzen ab. Die Patienten beurteilen ihren gegenwärtigen Zustand und fragen sich dann, wie sich die Dinge möglicherweise geändert haben könnten, um ein rückblickendes Urteil zu bilden [11]. In dieser prospektiv angelegten, klinischen Studie wurden 126 Patienten (Alter: 17 bis 83 Jahre, 49% Frauen), welche an der Poliklinik für Prothetik und Werkstoffkunde der Universität Leipzig behandelt wurden, einer Befragung unter Zuhilfenahme des OHIP-G49 unterzogen.

Die Patienten wurden zu Beginn der Therapie (Baseline, Pre-Test) und 4 bis 6 Wochen nach der Therapie (Follow-Up, Post-Test) befragt. Während der Follow-Up-Bewertung wurden die Patienten auch gebeten, ihren Mundgesundheitsstatus zu Beginn (Neubewertung, Then-Test) anhand eines speziell entworfenen Fragebogens rückwirkend zu bewerten. Die mundgesundheitsbezogene Lebensqualität (MLQ/OHRQoL) wird als einfache Summe aller 49 OHIP-Fragen mit Summenwerten zwischen 0 und 196 Punkten dargestellt. Höhere Werte deuten auf mehr Probleme

und damit auf eine niedrigere MLQ/ OHRQoL hin. Ebenso wurden die Patienten gebeten, ihre allgemeine Mundgesundheit und ihren allgemeinen Gesundheitszustand anhand einer fünfstufigen Skala zu bewerten („ausgezeichnet“ = 4, „sehr gut“ = 3, „gut“ = 2, „moderat“ = 1, „schlecht“ = 0). Der Respond Shift wurde als Differenz zwischen den OHIP-Summenwerten und dem anschließenden Then-Test berechnet. Darüber hinaus wurden die Patienten beim Follow-Up gebeten, die Änderung des Mundgesundheitsstatus anhand einer globalen Frage retrospektiv zu beurteilen. Die Bewertungen der Patienten zur Mundgesundheit und allgemeiner Gesundheit wurden als Validitätskriterien für die MLQ/ OHRQoL-Bewertungen herangezogen. Bei der statistischen Analyse wurde zuerst der Anteil der Patienten für jede Antwortkategorie der globalen Frage, bezüglich der wahrgenommenen oralen und allgemeinen Gesundheit in den drei Beurteilungszeiträumen (Baseline, Follow-Up und Then-Test), ermittelt. Anschließend wurden die Reliabilität und Validität von Baseline, Follow-Up und Then-Test verglichen. Dazu wurde die Interne Konsistenz als Maß für die Reliabilität durch die Berechnung von Cronbachs-Alpha [13] und durchschnittliche Inter-Item-Korrelation, bestimmt. Des Weiteren wurde untersucht, ob die Unterschiede zwischen der Baseline-Bewertung und dem Then-Test (Respond Shift) mit der Wahrnehmung der Patienten hinsichtlich Veränderungen der Mundgesundheit nach der Behandlung zusammenhängen. Daher berechneten wir den Mean-Wert und die entsprechenden 95% -Konfidenzintervalle (CI) für die Unterschiede in den OHIP-Summenwerten. Zusätzlich wurden für diese Kategorien standardisierte Effektgrößen (Cohens d) berechnet. Richtlinien legen nahe, dass eine Effektgröße von $d = 0,2$ klein; $0,5$ mittel und $0,8$ groß sind [14]. Schließlich wurden die Korrelationen zwischen den OHIP-Summenwerten der Baseline-Bewertung und des Then-Tests für die Kategorien der globalen Frage ermittelt und verglichen. Für die paarweisen Vergleiche wurden die Pearson-Produkt-Moment-Korrelationskoeffizienten berechnet. Wir stellten die Hypothese auf, dass die Korrelationen des Then-Tests mit dem tatsächlichen Ausgangswert und der Follow-Up-Bewertung unterschiedlich sind, insbesondere wenn die wahrgenommenen Veränderungen der Mundgesundheit nach der Behandlung groß sind. Wir interpretierten eine stärkere Korrelation des Then-Tests mit dem Follow-Up als Indikator für eine implizite Veränderungstheorie und eine stärkere Korrelation des Then-Tests mit der Baseline-Bewertung als Indikator für die kognitive Dissonanz. In den Ergebnissen zeigt sich, dass sich Baseline und retrospektiver Pre-Test (Then-

Test) hinsichtlich der internen Konsistenz und der Validität nicht wesentlich unterscheiden. Der Respond Shift war ausgeprägter, wenn die Patienten während der Behandlung eine große Veränderung der MLQ/ OHRQoL wahrnahmen. Außerdem korrelierte der Then-Test stärker mit dem Ausgangswert als mit der Follow-Up-Bewertung.

Nach Kenntnis der Autoren ist dies die erste Studie in der Zahnmedizin, die untersucht, ob Ausmaß und Richtung des Respond Shift mit den global wahrgenommenen Veränderungen verbunden sind, und ob der Then-Test entweder vom anfänglichen Baselinezustand, oder von den Wahrnehmungen bei der Nachuntersuchung abhängt.

Diese Ergebnisse sind von klinischer Bedeutung, da die von den Patienten nach der Therapie beurteilten Veränderungen der Mundgesundheit größer sind, als bei einer prospektiven Beurteilung. Eine Studie präferiert, sowohl eine prospektive als auch eine retrospektive Bewertung aufzunehmen, um Vergleiche und zu ermöglichen [15]. So zeigen unsere Ergebnisse, dass die retrospektive Bewertung von den Wahrnehmungen der Patienten beeinflusst wird. Dementsprechend sind die Ergebnisse der Behandlungseffekte aus prospektiven und retrospektiven Bewertungen nicht vergleichbar. In Anbetracht der Einschränkungen unserer Studie legen wir nahe, dass die kognitive Dissonanz bei der retrospektiven Beurteilung der mundgesundheitsbezogenen Lebensqualität unter Verwendung des OHIP wahrscheinlicher ist, als eine implizierte Veränderungstheorie. Weitere Studien sind jedoch erforderlich, um diese Hypothese zu belegen.

Referenzen

1. John, M.T., et al., *Oral health-related quality of life in Germany*. Eur J Oral Sci, 2003. **111**(6): p. 483-91.
2. John, M.T., D.L. Patrick, and G.D. Slade, *The German version of the Oral Health Impact Profile--translation and psychometric properties*. Eur J Oral Sci, 2002. **110**(6): p. 425-33.
3. Slade, G.D. and A.J. Spencer, *Development and evaluation of the Oral Health Impact Profile*. Community Dent Health, 1994. **11**(1): p. 3-11.
4. Allison, P., et al., *A cross-cultural study of oral health values*. J Dent Res, 1999. **78**(2): p. 643-9.

5. Reissmann, D.R., *Alignment of oral health-related with health-related quality of life assessment*. J Prosthodont Res, 2016. **60**(2): p. 69-71.
6. Allison, P.J., D. Locker, and J.S. Feine, *Quality of life: a dynamic construct*. Soc Sci Med, 1997. **45**(2): p. 221-30.
7. Sprangers, M.A. and C.E. Schwartz, *Integrating response shift into health-related quality of life research: a theoretical model*. Soc Sci Med, 1999. **48**(11): p. 1507-15.
8. Nieuwkerk, P.T., et al., *Are retrospective measures of change in quality of life more valid than prospective measures?* Med Care, 2007. **45**(3): p. 199-205.
9. McPhail, S. and T. Haines, *Response shift, recall bias and their effect on measuring change in health-related quality of life amongst older hospital patients*. Health Qual Life Outcomes, 2010. **8**: p. 65.
10. Festinger L, C.J., *Cognitive consequences of forced compliance*. J Abnorm Psychol, 1959. **58**: p. 203-210.
11. Norman, G., *Hi! How are you? Response shift, implicit theories and differing epistemologies*. Qual Life Res, 2003. **12**(3): p. 239-49.
12. Wilson, I.B., *Clinical understanding and clinical implications of response shift*. Soc Sci Med, 1999. **48**(11): p. 1577-88.
13. Bland, J.M. and D.G. Altman, *Cronbach's alpha*. BMJ, 1997. **314**(7080): p. 572.
14. Cohen, J., *A power primer*. Psychol Bull, 1992. **112**(1): p. 155-9.
15. Blome, C. and M. Augustin, *Measuring change in quality of life: bias in prospective and retrospective evaluation*. Value Health, 2015. **18**(1): p. 110-5.

Literaturverzeichnis

1. ARMELLINI DB, HEYDECKE G, WITTER DJ , CREUGERS NH. Effect of removable partial dentures on oral health-related quality of life in subjects with shortened dental arches: a 2-center cross-sectional study. *Int J Prosthodont* 2008; **21**: 524-530.
2. DE OLIVEIRA CM , SHEIHAM A. Orthodontic treatment and its impact on oral health-related quality of life in Brazilian adolescents. *J Orthod* 2004; **31**: 20-27; discussion 15.
3. HEYDECKE G, LOCKER D, AWAD MA, LUND JP , FEINE JS. Oral and general health-related quality of life with conventional and implant dentures. *Community Dent Oral Epidemiol* 2003; **31**: 161-168.
4. NICKENIG HJ, WICHMANN M, ANDREAS SK , EITNER S. Oral health-related quality of life in partially edentulous patients: assessments before and after implant therapy. *J Craniomaxillofac Surg* 2008; **36**: 477-480.
5. REISSMANN DR, SCHIERZ O, SZENTPETERY AG , JOHN MT. Improved perceived general health is observed with prosthodontic treatment. *J Dent* 2011; **Feb 16**: [Epub ahead of print].
6. SLADE GD. The Oral Health Impact Profile. In: SLADE GD, ed. *Measuring Oral Health and Quality of Life*. Chapel Hill: Department of Dental Ecology, School of Dentistry, University of North Carolina, 1997; 93-104.
7. ALLISON PJ, LOCKER D , FEINE JS. Quality of life: a dynamic construct. *Soc Sci Med* 1997; **45**: 221-230.
8. DALTRY LH, LARSON MG, EATON HM, PHILLIPS CB , LIANG MH. Discrepancies between self-reported and observed physical function in the elderly: the influence of response shift and other factors. *Soc Sci Med* 1999; **48**: 1549-1561.
9. MCPHAIL S , HAINES T. Response shift, recall bias and their effect on measuring change in health-related quality of life amongst older hospital patients. *Health Qual Life Outcomes* 2010; **8**: 65.
10. RAZMJOU H, SCHWARTZ CE, YEE A , FINKELSTEIN JA. Traditional assessment of health outcome following total knee arthroplasty was confounded by response shift phenomenon. *J Clin Epidemiol* 2009; **62**: 91-96.
11. SCHWARTZ CE , FINKELSTEIN JA. Understanding inconsistencies in patient-reported outcomes after spine treatment: response shift phenomena. *Spine J* 2009; **9**: 1039-1045.
12. SHI HY, LEE KT, LEE HH, UEN YH , CHIU CC. Response shift effect on gastrointestinal quality of life index after laparoscopic cholecystectomy. *Qual Life Res* 2010.
13. WILSON IB. Clinical understanding and clinical implications of response shift. *Soc Sci Med* 1999; **48**: 1577-1588.
14. SCHWARTZ CE, BODE R, REPUGGI N, BECKER J, SPRANGERS MA , FAYERS PM. The clinical significance of adaptation to changing health: a meta-analysis of response shift. *Qual Life Res* 2006; **15**: 1533-1550.
15. RING L, HOFER S, HEUSTON F, HARRIS D , O'BOYLE CA. Response shift masks the treatment impact on patient reported outcomes (PROs): the example of individual quality of life in edentulous patients. *Health Qual Life Outcomes* 2005; **3**: 55.

16. COHEN J. *Statistical power analysis for the behavioral sciences*, 2nd ed. Hillsdale, NJ: Lawrence Earlbaum Associates, 1988.
17. SLADE GD , SPENCER AJ. Development and evaluation of the Oral Health Impact Profile. *Community Dent Health* 1994; **11**: 3-11.
18. JOHN MT, HUJOEL P, MIGLIORETTI DL, LERESCHE L, KOEPSSELL TD , MICHEELIS W. Dimensions of oral-health-related quality of life. *J Dent Res* 2004; **83**: 956-960.
19. JOHN MT, LERESCHE L, KOEPSSELL TD, HUJOEL P, MIGLIORETTI DL , MICHEELIS W. Oral health-related quality of life in Germany. *Eur J Oral Sci* 2003; **111**: 483-491.
20. SCHWARTZ CE , SPRANGERS MA. Methodological approaches for assessing response shift in longitudinal health-related quality-of-life research. *Soc Sci Med* 1999; **48**: 1531-1548.
21. BROBERGER E, SPRANGERS M , TISHELMAN C. Do internal standards of quality of life change in lung cancer patients? *Nurs Res* 2006; **55**: 274-282.
22. OORT FJ. Using structural equation modeling to detect response shifts and true change. *Qual Life Res* 2005; **14**: 587-598.
23. VISSER MR, OORT FJ , SPRANGERS MA. Methods to detect response shift in quality of life data: a convergent validity study. *Qual Life Res* 2005; **14**: 629-639.
24. BORSBOOM GJ, KORFAGE IJ, ESSINK-BOT ML , DUIVENVOORDEN HJ. The structural equation modeling technique did not show a response shift, contrary to the results of the then test and the individualized approaches. *J Clin Epidemiol* 2007; **60**: 426-427; author reply 427-428.
25. JOHN MT, REISSMANN DR, SZENTPETERY A , STEELE J. An approach to define clinical significance in prosthodontics. *J Prosthodont* 2009; **18**: 455-460.
26. GUYATT GH, OSOBA D, WU AW, WYRWICH KW , NORMAN GR. Methods to explain the clinical significance of health status measures. *Mayo Clin Proc* 2002; **77**: 371-383.
27. NORMAN GR, SLOAN JA , WYRWICH KW. Interpretation of changes in health-related quality of life: the remarkable universality of half a standard deviation. *Med Care* 2003; **41**: 582-592.
28. NORMAN G. Hi! How are you? Response shift, implicit theories and differing epistemologies. *Quality of life research : an international journal of quality of life aspects of treatment, care and rehabilitation* 2003; **12**: 239-249.
29. FOLKMAN S , LAZARUS RS. The relationship between coping and emotion: implications for theory and research. *Soc Sci Med* 1988; **26**: 309-317.
30. LAZARUS RS , FOLKMAN S. *Stress, appraisal and coping*, 1st ed. New York: Springer Publishing Company, 1984.
31. FOLKMAN S. Personal control and stress and coping processes: a theoretical analysis. *J Pers Soc Psychol* 1984; **46**: 839-852.
32. HEDLUND M, RONNE-ENGSTROM E, CARLSSON M , EKSELIUS L. Coping strategies, health-related quality of life and psychiatric history in patients with aneurysmal subarachnoid haemorrhage. *Acta Neurochir (Wien)* 2010; **152**: 1375-1382; discussion 1382.
33. YANG HC, BROTHERS BM , ANDERSEN BL. Stress and quality of life in breast cancer recurrence: moderation or mediation of coping? *Ann Behav Med* 2008; **35**: 188-197.

34. DEMPSTER M, CARNEY R , MCCLEMENTS R. Response shift in the assessment of quality of life among people attending cardiac rehabilitation. *Br J Health Psychol* 2010; **15**: 307-319.
35. PUCHEU S, CONSOLI SM, D'AUZAC C, FRANCAIS P , ISSAD B. Do health causal attributions and coping strategies act as moderators of quality of life in peritoneal dialysis patients? *J Psychosom Res* 2004; **56**: 317-322.
36. TAYLOR JL, SMITH PJ, BABYAK MA, BARBOUR KA, HOFFMAN BM, SEBRING DL, DAVIS RD, PALMER SM, KEEFE FJ, CARNEY RM, CSIK I, FREEDLAND KE , BLUMENTHAL JA. Coping and quality of life in patients awaiting lung transplantation. *J Psychosom Res* 2008; **65**: 71-79.
37. FESTINGER L. A theory of social comparison processes. *Human Relations* 1954; **7**: 117-140.
38. FESTINGER L , CARLSMITH JM. Cognitive consequences of forced compliance. *Journal of abnormal psychology* 1959; **58**: 203-210.
39. SCIENTIFIC ADVISORY COMMITTEE OF THE MEDICAL OUTCOMES TRUST. Assessing health status and quality-of-life instruments: attributes and review criteria. *Qual Life Res* 2002; **11**: 193-205.
40. KIEFFER JM , HOOGSTRATEN J. Item-order effects in the Oral Health Impact Profile (OHIP). *Eur J Oral Sci* 2008; **116**: 245-249.
41. JOHN MT, REISSMANN DR, SCHIERZ O , ALLEN F. No significant retest effects in oral health-related quality of life assessment using the Oral Health Impact Profile. *Acta Odontol Scand* 2008; **66**: 135-138.
42. REISSMANN DR, JOHN MT , SCHIERZ O. Influence of administration method on oral health-related quality of life assessment using the Oral Health Impact Profile. *Eur J Oral Sci* 2011; **119**: 73-78.
43. KIEFFER JM, VERRIPS E , HOOGSTRATEN J. Model specification in oral health-related quality of life research. *Eur J Oral Sci* 2009; **117**: 481-484.
44. HASSEL AJ, ROLKO C, KOKE U, LEISEN J , RAMMELSBERG P. A German version of the GOHAI. *Community Dent Oral Epidemiol* 2008; **36**: 34-42.
45. MCGRATH C, COMFORT MB, LO EC , LUO Y. Patient-centred outcome measures in oral surgery: validity and sensitivity. *Br J Oral Maxillofac Surg* 2003; **41**: 43-47.
46. KORFAGE IJ, DE KONING HJ , ESSINK-BOT ML. Response shift due to diagnosis and primary treatment of localized prostate cancer: a then-test and a vignette study. *Qual Life Res* 2007; **16**: 1627-1634.

Anlagen

Erklärung des eigenen Beitrags an der Publikation

Mein Beitrag zur vorliegenden Studie umfasste zunächst eine umfassende Literaturrecherche unter anderem online in der Literaturlatenbank PubMed, aber auch eine Handsuche in relevanten Journalen und in Referenzlisten von identifizierten Artikeln. Anschließend wurde mit Hilfe von PD Dr. Reißmann das Ziel der Studie definiert, eine konkrete Fragestellung entwickelt und eine Studienprotokoll erstellt. Die Erstellung der Studienunterlagen und Fragebögen erfolgte ebenfalls durch mich.

Ich führte die Befragung der Patienten im Studentenkurs der Poliklinik für Prothetik und Werkstoffkunde der Universität Leipzig selbstständig durch. Dies beinhaltete die Auswahl der Patienten, deren Aufklärung und das Einholen der Einverständniserklärung. Während des Ausfüllens der Fragebögen wurden die Patienten von mir unterstützt und betreut. Die Organisation der Nachkontrolltermine, die Durchführung der Nachkontrollen und die Betreuung des Ausfüllens der Fragebögen wurden ebenfalls von mir selbstständig durchgeführt.

Des Weiteren war ich für die Erfassung und Übertragung der Fragebögen in eine Datenbank zuständig. Dazu wurden die Daten von mir digitalisiert und einer Qualitätskontrolle sowie einem Plausibilitätscheck unterzogen.

Bei der Auswahl geeigneter statistischer Analysen und deren Auswertung erhielt ich praktische und fachliche Unterstützung von PD Dr. Reißmann.

Ich erstellte eine Rohfassung des Manuskripts mit Teilen der Einleitung, der Methodik, der Ergebnisse und Diskussion. Die Darstellung der statistischen Auswertung wurde hauptsächlich von PD Dr. Reißmann erstellt, welcher auch die erste Version des Manuskripts überarbeitete und mir Hilfestellungen für die Finalisierung des Manuskripts gab. Gemeinsam brachten wir die Ergebnisse in einen wissenschaftlichen Zusammenhang und erörterten deren klinische Relevanz. Darüber hinaus war ich an der Einarbeitung der Gutachterkommentare und der Erstellung der finalen Version des Manuskripts beteiligt.

02.07.18

Datum, Unterschrift Doktorand
Antje Erler

Hiermit bestätige ich die Angaben der Doktorandin.

28.06.18
PD Dr. Daniel R. Reißmann
Oberarzt

Datum, Unterschrift korrespondierender Autor
PD Dr. D. Reißmann

13.07.18

Datum, Unterschrift Mitautor
Prof. Dr. C. Hirsch

Prof. Dr. Christian Hirsch, MSc
Direktor der Poliklinik für
Kinderzahnheilkunde und
Primärprävention

02.08.2018

Datum, Unterschrift Mitautor
Dr. O. Schierz

Erklärung über die eigenständige Abfassung der Arbeit

Hiermit erkläre ich, dass ich die vorliegende Arbeit selbstständig und ohne unzulässige Hilfe oder Benutzung anderer als der angegebenen Hilfsmittel angefertigt habe. Ich versichere, dass Dritte von mir weder unmittelbar noch mittelbar eine Vergütung oder geldwerte Leistungen für Arbeiten erhalten haben, die im Zusammenhang mit dem Inhalt der vorgelegten Dissertation stehen, und dass die vorgelegte Arbeit weder im Inland noch im Ausland in gleicher oder ähnlicher Form einer anderen Prüfungsbehörde zum Zweck einer Promotion oder eines anderen Prüfungsverfahrens vorgelegt wurde. Alles aus anderen Quellen und von anderen Personen übernommene Material, das in der Arbeit verwendet wurde oder auf das direkt Bezug genommen wird, wurde als solches kenntlich gemacht. Insbesondere wurden alle Personen genannt, die direkt an der Entstehung der vorliegenden Arbeit beteiligt waren. Die aktuellen gesetzlichen Vorgaben in Bezug auf die Zulassung der klinischen Studien, die Bestimmungen des Tierschutzgesetzes, die Bestimmungen des Gentechnikgesetzes und die allgemeinen Datenschutzbestimmungen wurden eingehalten. Ich versichere, dass ich die Regelungen der Satzung der Universität Leipzig zur Sicherung guter wissenschaftlicher Praxis kenne und eingehalten habe.

24.11.2018
Datum

Unterschrift

Lebenslauf

Persönliche Daten

Name
Geburtsdatum
Geburtsort
Familienstand

Antje Erler geb. Remmler
22.07.1982
Werdau
verheiratet seit 30.12.2015

Schulische Ausbildung 1993 – 2001

Julius-Motteler-Gymnasium Crimmitschau, Abitur

Berufliche Ausbildung 09/ 2001 – 08/ 2002

VW - Bildungsinstitut Zwickau,
Berufsvorbereitendes Jahr (Wirtschaft,
Informatik)
Dehner GmbH & Co.KG, Rain am Lech,
Duales Studium zum Handelsfachwirt

09/ 2002 – 09/ 2004

Hochschulausbildung 10/ 2004 – 10/ 2009

Studium der Zahnmedizin an der Universität
Leipzig
Staatsexamen, Gesamtnote: gut

28.09.2009

Publikationen 12/ 2012

"Impact of response shift on assessment of
treatment effects using the Oral Health Impact
Profile", European
Journal of Oral Sciences

03/ 2018

„Bias in retrospective assessment of perceived
dental treatment effects when using the Oral
Health Impact Profile.“, Quality of Life Research

Berufliche Tätigkeit 01/ 2010 – 10/ 2011

Assistenz Zahnärztin in der Zahnarztpraxis
Dr. Marion Quandt, Markkleeberg

11/ 2011 – 08/ 2012

Weiterbildungsassistentin in der Praxis
Dr. Peter Ludwig, Gera

08/ 2012 – 07/ 2014

Elternzeit

07/ 2014 – 12/ 2015

Weiterbildungsassistentin in der Praxis
Dres. K. und A. Höhle, Radebeul

01/ 2016 – 04/ 2017

Beschäftigungsverbot, Elternzeit ab 06. Juli 2016

seit 05/ 2017

Weiterbildungsassistentin in der Praxis
Dr. Andreas Höhle, Radebeul

Meißen, 24.11.2018

Antje Erler

Publikationen

Eur J Oral Sci. 2012 Dec;120(6):520-5. doi: 10.1111/j.1600-0722.2012.00998.x. Epub 2012 Oct 8.

Impact of response shift on the assessment of treatment effects using the Oral Health Impact Profile.

Reissmann DR¹, Remmler A, John MT, Schierz O, Hirsch C.

Qual Life Res. 2018 Mar;27(3):775-782. doi: 10.1007/s11136-017-1725-z. Epub 2017 Oct 23.

Bias in retrospective assessment of perceived dental treatment effects when using the Oral Health Impact Profile.

Reissmann DR¹, Erler A², Hirsch C³, Sierwald I⁴, Machuca C⁵, Schierz O².

1

Department of Prosthetic Dentistry, Center for Dental and Oral Medicine, University Medical Center Hamburg-Eppendorf, Martinistrasse 52, 20246, Hamburg, Germany.
d.reissmann@uke.de.

2

Department of Prosthodontics and Materials Science, School of Dentistry, University of Leipzig, Leipzig, Germany.

3

Department of Pediatric Dentistry, School of Dentistry, University of Leipzig, Leipzig, Germany.

4

Department of Orthodontics, Dentofacial Orthopedics and Pedodontics, Center for Dental and Craniofacial Sciences, Charité - Universitätsmedizin Berlin, Berlin, Germany.

5

Academic Unit of Dental Public Health, School of Clinical Dentistry, University of Sheffield, Sheffield, UK.

Danksagung

Ohne die Unterstützung zahlreicher Personen und Institutionen hätte meine Publikationspromotion in dieser Form nicht realisiert werden können. Für die vielfältige Hilfe möchte ich mich an dieser Stelle sehr herzlich bedanken.

Mein besonderer Dank gilt zunächst meinem Doktorvater Prof. Dr. Christian Hirsch und meinem Betreuer Dr. Daniel R. Reißmann, die meine Arbeit stets mit viel Verständnis, Geduld und konstruktiver Kritik unterstützt haben.

Für die vielfältige Unterstützung bin ich besonders Frau Annett Schrock dankbar, die mir bei der Erstellung der statistischen Analysen, neben Herrn Dr. Reißmann, eine große Hilfe war.

Mein besonderer Dank gilt schließlich meiner lieben Studienfreundin, Melanie Schädlich, die mich bei der Datensammlung und Nachuntersuchung der Patienten unterstützt hat.

Der größte Dank gebührt meinem Mann, Michael Erler, ohne dessen Unterstützung die vorliegende Arbeit nicht möglich gewesen wäre. Ihm und unseren Kindern, Friedrich und Leonora, ist diese Studie gewidmet.