

# Die Anwendung Hierarchischer Linearer Modelle für Einrichtungsvergleiche in der Qualitätssicherung und Rehabilitationsforschung

E. Farin

*Using Hierarchical Linear Modelling for Rehabilitation Centre Comparisons in Quality Assurance and Rehabilitation Research*

## Zusammenfassung

Ein fairer Vergleich von Rehabilitationskliniken im Hinblick auf Qualität und Effektivität (z. B. im Rahmen von Qualitätssicherungsprogrammen) setzt voraus, dass die relevanten Einflussfaktoren des Rehabilitationserfolgs, die von der Rehabilitationseinrichtung nicht beeinflussbar sind (die sog. „Confounder“ wie z. B. Komorbidität und Alter des Patienten bei Reha-Beginn), statistisch kontrolliert werden. Häufig werden dazu einfache lineare Regressionsanalysen ohne Zufallseffekte und ohne Interaktionsterme eingesetzt. Dieses Verfahren weist aber einige Beschränkungen auf, die sich durch die Methode der Hierarchischen Linearen Modelle (HLMs) beheben lassen. Die Bezeichnung „HLM“ drückt aus, dass für die verschiedenen hierarchisch gestuften Datenebenen einer Fragestellung (im Beispiel der Rehabilitationseinrichtungen z. B. die Ebene der Patienten und die Ebene der Kliniken) separate lineare Modelle aufgestellt werden. Dadurch können auch Einflussfaktoren auf der Ebene der Kliniken (z. B. das mittlere Alter der Patienten einer Einrichtung) berücksichtigt werden. Darüber hinaus kann mit den HLMs empirisch geprüft werden, ob sich die Regressionskoeffizienten (also z. B. der Einfluss des Alters auf das Rehabilitationsergebnis) zwischen den Einrichtungen bedeutsam unterscheiden. Somit weisen HLMs gegenüber gewöhnlichen regressionsanalytischen Verfahren die Vorteile auf, dass sie die Multiebenenstruktur des Vergleichsproblems berücksichtigen, die Aufnahme von Prädiktoren auf der Ebene der Einrichtungen ermöglichen und die Modellie-

## Abstract

For a fair comparison of rehabilitation centres with respect to the effects of the treatment provided (e.g. for the purpose of quality assurance programmes), it is essential that those factors which influence the outcome of rehabilitation treatment and over which the rehabilitation centres have no control (the so-called „confounders“, such as co-morbidity and age of the patients on commencement of treatment) are included in the statistical analysis. Simple linear regression models without random effects and without interaction terms are frequently used for this purpose. However, this method has certain limitations which can be avoided if hierarchical linear modelling (HLM) is employed. HLM has the advantage over standard regression analysis methods in that it can be used to take into account the multi-level structure of a comparison problem, allows predictors to be introduced at the level of the centres and also makes it possible to model variations of regression coefficients for the centres. When the HLM technique is used, separate linear models can be produced for the various hierarchically structured data levels of the question (e.g. the levels „patients“ and „centres“ for rehabilitation centres, for example). Moreover, it can be empirically tested with HLMs whether the rehabilitation coefficients (e.g. effects of mean age of patients on the outcome of rehabilitation) differ significantly between the centres. In this article, we describe the use of hierarchical linear modelling on the basis of data obtained from the quality assurance programme of the statutory

### Koordinatoren der Reihe „Methoden in der Rehabilitationsforschung“:

Prof. Dr. Dr. Hermann Faller, Würzburg; Prof. Dr. Thomas Kohlmann, Greifswald;  
Dr. Christian Zwingmann, Berlin  
Interessenten, die einen Beitrag zur Reihe beisteuern möchten, werden gebeten,  
vorab Kontakt aufzunehmen, E-mail: christian.zwingmann@vdr.de

### Institutsangaben

Universitätsklinikum Freiburg, Abt. Qualitätsmanagement und Sozialmedizin

### Korrespondenzadresse

Dr. phil. Dipl.-Psych. Erik Farin · Universitätsklinikum Freiburg · Abt. Qualitätsmanagement und Sozialmedizin  
· Breisacherstraße 62, Haus 4 · 79106 Freiburg · E-mail: erik.farin@uniklinik-freiburg.de

### Bibliografie

Rehabilitation 2005; 44: 157 – 164 © Georg Thieme Verlag KG Stuttgart · New York  
DOI 10.1055/s-2004-834785  
ISSN 0034-3536

rung der Variation von Regressionskoeffizienten über die Einrichtungen erlauben. Im Beitrag wird die Anwendung der Hierarchischen Linearen Modelle anhand von Daten aus dem Qualitätssicherungsprogramm der gesetzlichen Krankenkassen in der medizinischen Rehabilitation (QS-Reha-Verfahren) verdeutlicht.

### Schlüsselwörter

Hierarchische Lineare Modelle · Multiebenenanalyse · Risikoadjustierung · Qualitätssicherung

health insurance schemes in the field of medical rehabilitation („QS-Reha“).

### Key words

Hierarchical linear models · multilevel analysis · risk adjustment · quality management

## Einleitung

Der Vergleich von Einrichtungen (Rehabilitationskliniken, Krankenhäusern, ambulanten Reha-Zentren etc.) im Hinblick auf Qualität und Effektivität stellt einen zentralen Bestandteil von Qualitätssicherungsprogrammen dar, den auch der Gesetzgeber im Sozialgesetzbuch (SGB) fordert (vgl. SGB IX §20). Aber wie kann man auf faire Weise Rehabilitationskliniken in Qualitätssicherungsverfahren oder rehabilitationswissenschaftlichen Studien vergleichen? Das Grundproblem besteht darin, dass der jeweils betrachtete Outcomeparameter (hier und im Folgenden wird als Beispiel der funktionale Status bei Entlassung als Beispiel herangezogen) nicht nur von der Qualität der Einrichtung abhängt, sondern neben einem Zufallsfaktor bzw. Messfehler auch von Patienteneigenschaften (z.B. Eingangsstatus bez. des jeweils betrachteten Outcomeparameters, Komorbidität, Alter). Da im Rahmen einer routinemäßigen Qualitätssicherung keine kontrollierte Studie mit randomisierter Zuweisung erfolgen kann, sondern in der Regel Beobachtungsstudien durchgeführt werden (vgl. z.B. [1]), wird die Verteilung dieser sog. „Confounder“ in verschiedenen Einrichtungen unterschiedlich ausfallen, wodurch die Vergleichbarkeit dieser Einrichtungen infrage gestellt ist. Ein „naiver“ Vergleich der bei den Kliniken im Mittel bei Entlassung erzielten Werte des funktionalen Status ist ebenso inadäquat wie ein Vergleich der Prä-Post-Effektstärken, da dieses Maß lediglich den Eingangsstatus, nicht aber weitere relevante Confounder berücksichtigt.

Ein elaborierteres und mittlerweile häufig eingesetztes Vorgehen besteht darin, mithilfe einer linearen Regressionsanalyse das Kriterium „Funktionaler Status bei Entlassung“ vorherzusagen. Als Prädiktoren gehen in diese Regressionsanalyse die gemäß vorliegender empirischer Studien (vgl. [2] für eine aktuelle Übersicht) relevanten patientenbezogenen Einflussfaktoren des Rehabilitationserfolgs ein. Wichtig ist dabei jedoch, dass nur solche Variablen Berücksichtigung finden, die von der Klinik nicht beeinflussbar sind. Das sind in der Regel Patientencharakteristika zum Aufnahmezeitpunkt, wie z.B. der funktionale und psychische Status, deren Werte zwar im Laufe der Behandlung verändert werden können, deren Aufnahmewerte aber von der Klinik nicht beeinflusst werden können, da auf die Patientenzuweisung nicht oder nur sehr begrenzt eingewirkt werden kann. Die von der Klinik beeinflussbaren, also durch qualitätsrelevantes Handeln veränderbaren Faktoren (z.B. Personalschlüssel oder Reha-Motivation nach einer Behandlungswoche) dürfen hingegen nicht kontrolliert werden. Die diesbezügliche Varianz, welche Qualitätsunterschiede der Einrichtungen abbildet, soll durch das Verfahren ja gerade aufgezeigt werden.

Die Residualwerte der Regressionsanalyse (d.h. die Differenzen zwischen den tatsächlichen und den durch die Regression vorhergesagten Werte) stellen einen vom Einfluss der jeweils berücksichtigten Prädiktoren befreiten Outcomeparameter dar, der für aussagekräftige Einrichtungsvergleiche herangezogen werden kann (vgl. z.B. [10,12]). Ist es möglich, ein Regressionsmodell an einer sehr großen, für die zu vergleichenden Einrichtungen repräsentativen Stichprobe zu bestimmen, so wäre es denkbar, Residualwerte für neu zu analysierende Kliniken mithilfe der einmal bestimmten Regressionsgleichung zu errechnen. Ein genaueres Vorgehen besteht jedoch darin, mit den Daten der neu in die Analyse aufzunehmenden Einrichtungen ein aktualisiertes Regressionsmodell zu berechnen. Über weitere konzeptionelle Fragen im Rahmen von regressionsanalytischen „Risikoadjustierungsverfahren“<sup>1</sup> informiert z.B. das von Iezzoni herausgegebene Standardwerk [3].

„Einfache“ lineare Regressionsanalysen (d.h. multiple lineare Regressionsmodelle ohne zufällige Effekte und ohne Wechselwirkungsterme) oder statistisch weitgehend äquivalente Kovarianzanalysen mit der Einrichtung als unabhängiger Variable, dem jeweils betrachteten Outcomeparameter als abhängiger Variable und den Confoundern als Kovariaten sind jedoch nicht geeignet, zwei Probleme zu lösen, die sich bei näherer Betrachtung ergeben:

1. Wie kann man den Umstand berücksichtigen, dass die Regressionskoeffizienten (Achsenabschnitt und Steigungskoeffizienten bzw. – im Englischen – „Intercepts“ und „Slopes“, s. Abb. 1) der zur Adjustierung herangezogenen Regressionsgleichung unter Umständen je nach Rehabilitationsklinik variieren (z.B. könnte es sein, dass der negative Einfluss der psychischen Komorbidität auf das zu erwartende Reha-Outcome in Kliniken mit einem interdisziplinären Therapiekonzept geringer ausfällt)?
2. Wie soll man damit umgehen, dass es auch auf der Ebene der Einrichtungen (nicht nur auf der Ebene der Patienten) Merkmale geben wird, die einen Einfluss auf das Outcome besitzen, von der Einrichtung nicht beeinflussbar sind und insofern bei Vergleichen adjustiert werden sollten (z.B. könnte es sein, dass das mittlere Alter der Patienten einer Rehabilitationsklinik im Sinne eines Kontextfaktors einen Einfluss auf das individuelle Reha-Outcome besitzt)?

<sup>1</sup> Der Begriff „Risikoadjustierungsverfahren“ bringt zum Ausdruck, dass nach den unterschiedlichen „Risiken“ für einen geringen Behandlungserfolg (jenes Risiko ist z.B. bei einem wenig motivierten, multimorbiden Patienten höher) adjustiert wird, um faire Vergleiche zu ermöglichen.

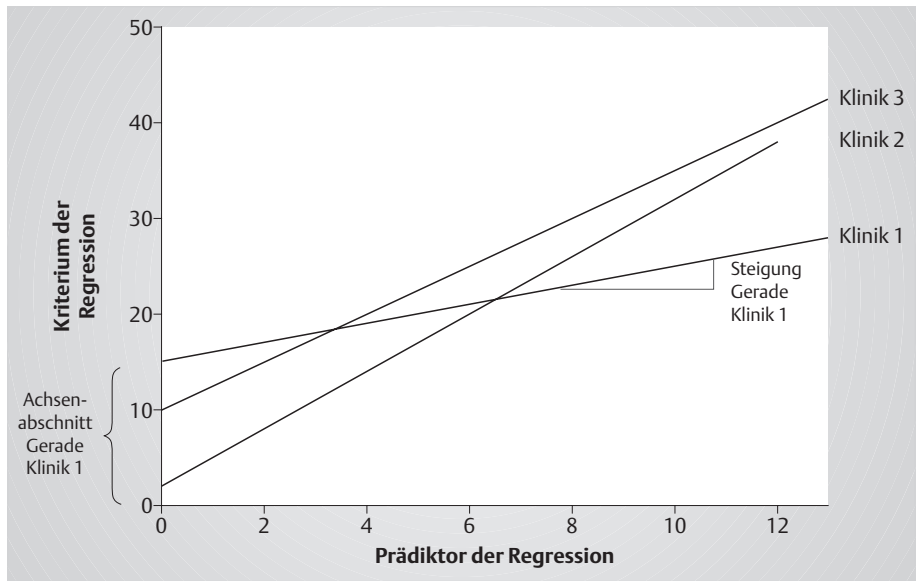


Abb. 1 Die Regressionsgeraden von drei Kliniken mit verschiedenen Steigungskoeffizienten („slopes“) und Achsenabschnitten („intercepts“).

Beide Probleme lassen sich auf methodisch korrekte Weise mit einfachen linearen Regressionsanalysen nicht bearbeiten, sie sind zu lösen mithilfe der Anwendung so genannter „Hierarchischer Linearer Modelle“ (kurz: HLMs), die die Multiebenenstruktur der Daten adäquat abbilden: Patienten (Ebene 1) sind in Einrichtungen (Ebene 2) gruppiert. Die folgende Darstellung der HLMs dürfte in ihren Grundzügen auch für Leser ohne besondere statistische Vorkenntnisse verständlich sein, das Verständnis der Formeln erfordert jedoch teilweise die Kenntnis der mathematischen Grundlagen von einfachen linearen Regressionsmodellen.

**Die Grundlagen Hierarchischer Linearer Modelle**

Eine umfassende, aber methodisch relativ anspruchsvolle Einführung zu Hierarchischen Linearen Modellen liegt mit dem Buch von Raudenbush u. Bryk [4] vor; für „Neueinsteiger“ leichter zu lesen ist die Einführung von Kreft u. de Leeuw [5]. Die praktische Berechnung von Hierarchischen Linearen Modellen ist mit Standardstatistiksoftware möglich (z.B. in SPSS mit der Prozedur MIXED und in SAS mit der Prozedur PROC MIXED); konkrete Hinweise dazu gibt Singer [6]. In der Benutzeroberfläche aber häufig komfortabler sind verfahrensspezifische Softwarelösungen wie z.B. HLM [7] und MLwiN [8]. Zudem integrieren die verfahrensspezifischen Softwarelösungen wie z.B. HLM die von verschiedener Seite (z.B. [9]) geforderten Empirical-Bayes-Schätzer, die die Klinikmittelwerte zusätzlich hinsichtlich zufälliger Werte korrigieren.

Das Prinzip der Hierarchischen Linearen Modelle besteht darin, dass sie für die verschiedenen Ebenen der Betrachtung (im Folgenden werden der Einfachheit halber immer nur zwei Ebenen angenommen: die der Patienten und die der Klinik; weitere Ebenen wie „Stationen“ oder „Abteilungen“ wären aber möglich) getrennte lineare Modelle aufstellen und diese für die Anwendung der Parameterschätzverfahren durch Einsetzung der entsprechenden Modellgleichungen miteinander verknüpfen. Da die gewöhnliche Kovarianzanalyse bzw. lineare Regressionsanalyse einen Spezialfall der Hierarchischen Linearen Modelle darstellt,

kann man mit den Formeln der HLMs auch diese einfacheren Verfahren abbilden. Wir nehmen dazu den Fall an, dass das Outcome „Funktionaler Status bei Entlassung“ (Ft1) mittels der Prädiktoren „Funktionaler Status bei Aufnahme“ (Ft0) und Alter (Alter) vorhergesagt werden soll, um ein vom Einfluss der Confounder befreites Residuum zu erhalten. Die Formeln des HLM lauten dann:

Ebene 1 (Patienten):

$$Ft1_{ij} = \beta_{0j} + \beta_{1j} \times Ft0_{ij} + \beta_{2j} \times Alter_{ij} + r_{ij} \tag{Gl. 1}$$

(hier und in den folgenden Formeln gilt: i ist die Laufvariable bezüglich der Patienten, j ist die Laufvariable bezüglich der Einrichtungen)

Ebene 2 (Kliniken):

$$\beta_{0j} = \gamma_{00} \tag{Gl. 2 a}$$

$$\beta_{1j} = \gamma_{10} \tag{Gl. 2 b}$$

$$\beta_{2j} = \gamma_{20} \tag{Gl. 2 c}$$

$\beta_{0j}$ ,  $\beta_{1j}$  und  $\beta_{2j}$  sind die Regressionskoeffizienten,  $r_{ij}$  ist ein Fehlerterm, der den spezifischen Effekt bei einem einzelnen Patienten abbildet und für den Normalverteilung angenommen wird. i ist die Laufvariable der Personen und j – als Laufvariable der Einrichtungen – deutet an, dass die drei Regressionskoeffizienten im Prinzip je nach Einrichtung unterschiedliche Werte annehmen können. Im Spezialfall der üblichen linearen Regression ist das aber nicht der Fall: Die Gleichungen 2 a – 2 c (die diesen Spezialfall abbilden) zeigen, dass die drei Regressionskoeffizienten des Modells (das Intercept  $\beta_0$ , der Regressionskoeffizient des funktionalen Status bei Aufnahme  $\beta_1$  und der Regressionskoeffizient des Alters  $\beta_2$ ) auf der Ebene der Kliniken konstante Werte annehmen, also nicht variieren. Die konstanten Werte dieser drei Regressionskoeffizienten werden mit  $\gamma_{00}$ ,  $\gamma_{10}$  und  $\gamma_{20}$  bezeichnet.

Soll ein allgemeineres Modell aufgestellt werden, um das oben aufgeführte Problem 1 (die Variation der Regressionskoeffizienten) zu lösen, so werden in den Gleichungen auf der Ebene 2 Fehlerterme zugelassen, die den spezifischen Einfluss einer Einrichtung auf die Regressionskoeffizienten abbilden:

$$\beta_{0j} = \gamma_{00} + u_{0j} \quad (\text{Gl. 3 a})$$

$$\beta_{1j} = \gamma_{10} + u_{1j} \quad (\text{Gl. 3 b})$$

$$\beta_{2j} = \gamma_{20} + u_{2j} \quad (\text{Gl. 3 c})$$

Dieses so definierte HLM wird in der Literatur (vgl. z. B. [4]) als „Random-Coefficient Regression Model“ bezeichnet. Die Regressionskoeffizienten einer Einrichtung  $j$  (also hier im Beispiel das Intercept  $\beta_{0j}$  sowie die Steigungen bez. „Funktionaler Status bei Aufnahme“ und Alter) setzen sich zusammen aus einem konstanten Wert  $\gamma$  und einem für die Einrichtung  $j$  spezifischen Term  $u$ .  $u_{0j}$ ,  $u_{1j}$  und  $u_{2j}$  sind zufällige Effekte, die eine Variation der Regressionskoeffizienten über die Einrichtungen hinweg modellieren. Das bedeutet, dass die Koeffizienten in der Regel bei jeder Einrichtung einen anderen Wert annehmen. Im oben dargestellten Modell ohne Variation der Koeffizienten (Gl. 2 a–2 c) wurde hingegen für alle Einrichtungen derselbe Regressionskoeffizient angenommen.

Um auch Prädiktoren der Ebene 2 (also organisationale Variablen, siehe das oben erwähnte Problem 2) zuzulassen, werden in den Gleichungen zusätzliche Einflussfaktoren aufgenommen. Angenommen, es soll ein Einfluss des mittleren Alters der Patienten einer Klinik  $j$  ( $\text{AlterM}_j$ ) a) auf das Intercept und b) auf den Steigungskoeffizienten des funktionalen Status bei Aufnahme abgebildet werden, so lauten die Gleichungen:

$$\beta_{0j} = \gamma_{00} + \gamma_{01} \times \text{AlterM}_j + u_{0j} \quad (\text{Gl. 4 a})$$

$$\beta_{1j} = \gamma_{10} + \gamma_{11} \times \text{AlterM}_j + u_{1j} \quad (\text{Gl. 4 b})$$

$$\beta_{2j} = \gamma_{20} + u_{2j} \quad (\text{Gl. 4 c})$$

Inhaltlich besagt dies, dass a) das mittlere Alter in einer Klinik das Niveau des Outcomeparameters funktionaler Status bei Entlassung beeinflusst und b) dass das mittlere Alter den Einfluss des Eingangswerts funktionaler Status auf den Entlasswert beeinflusst.

Setzt man die Gleichungen 4a–4c in die Gleichung 1 ein, so erhält man das komplette Modell eines so genannten „Intercepts-and Slopes-as-Outcomes“-Ansatzes ([4], S. 80 ff.):

$$\begin{aligned} \text{Ft1}_{ij} &= \gamma_{00} + \gamma_{01} \times \text{AlterM}_j + u_{0j} + (\gamma_{10} + \gamma_{11} \times \text{AlterM}_j + u_{1j}) \\ &\times \text{Ft0}_{ij} + (\gamma_{20} + u_{2j}) \times \text{Alter}_{ij} + \epsilon_{ij} \end{aligned} \quad (\text{Gl. 5})$$

Das in Gl. 5 abgebildete Modell wird auch als „gemischtes lineares Modell“ bezeichnet, da in ihm zufällige Effekte und feste Effekte kombiniert werden. Die Gleichung besagt, dass sich der funktionale Status bei Entlassung eines Patienten  $i$  in der Einrichtung  $j$  ( $\text{Ft1}_{ij}$ ) berechnen lässt über

- ein Intercept (der sich wiederum aus einem konstanten Anteil  $\gamma_{00}$ , einem durch das mittlere Patientenalter in der Einrich-

tung bestimmten Anteil  $\gamma_{01} \times \text{AlterM}_j$  und aus einem einrichtungsspezifischen Faktor  $u_{0j}$  zusammensetzt),

- einen durch den Eingangswert des funktionalen Status bestimmten Anteil (dessen Höhe auch abhängt vom mittleren Patientenalter und einem einrichtungsspezifischen Faktor) und über
- einen durch das Alter des Patienten bestimmten Anteil (dessen Höhe auch abhängt von einem einrichtungsspezifischen Faktor).

### Drei typische Problemstellungen

Welche Variante der Hierarchischen Linearen Modelle zu verwenden ist, hängt von der Fragestellung ab, die im Rahmen eines Forschungs- oder Qualitätssicherungsprojekts verfolgt wird. Auf drei typische Problemstellungen soll im Folgenden eingegangen werden:

1. Welches sind die generellen Einflussfaktoren des Rehabilitationserfolgs? (die Prädiktorfragestellung)
2. Welche Einrichtung ist global betrachtet (über alle Patienten hinweg) im Vergleich zu anderen Kliniken besonders gut bzw. nicht so gut? (die Benchmarkingfragestellung)
3. In welcher Klinik ist für einen bestimmten Patienten der beste Rehabilitationserfolg zu erwarten? (die Zuweisungsfragestellung)

#### 1. Die Prädiktorfragestellung

Sucht man nach allgemeinen Einflussfaktoren des Rehabilitationserfolgs auf der patientenbezogenen *und* der organisationalen Ebene, so ist ein HLM anzuwenden, welches (z. B. gemäß Gl. 4a und 4b) sowohl Ebene-1- als auch Ebene-2-Prädiktoren enthält. Würde man die Ebene-2-Prädiktoren (z. B. mittleres Alter der Patienten einer Klinik) lediglich als zusätzliche Variablen auf der Patientenebene in den Datensatz einfügen und auf der Ebene 1 eine einfache lineare Regression mit dem mittleren Alter als zusätzlichen Prädiktor rechnen, so würde man zu verzerrten Schätzungen und Teststatistiken gelangen. Der Standardfehler des mittleren Alters würde unterschätzt werden, da auf der Ebene 1 jeder Patient berücksichtigt wird und das mittlere Alter der Patienten einer Einrichtung als einrichtungskonstanter Wert  $n$ -mal pro Klinik ( $n$  = Anzahl der Patienten pro Klinik) in die Berechnung eingehen würde.

Bei Studien außerhalb von Klinikvergleichen kann man in die HLMs auch organisationale Variablen aufnehmen, die von der Klinik beeinflusst werden können und Qualitätsaspekte abbilden (z. B. Personalschlüssel, Betriebsklima), um die Relevanz dieser einrichtungsbezogenen Kontextfaktoren für das Rehabilitationsergebnis zu prüfen. Ob man zusätzlich zur Aufnahme der Ebene-2-Prädiktoren eine einrichtungsbezogene Variation der Koeffizienten (dargestellt durch den Term  $u_{ij}$ , vgl. Gl. 4a und 4b) annimmt oder ob man eine Konstanz der Regressionskoeffizienten über die Einrichtungen postuliert, hängt von der jeweiligen Fragestellung ab. Die Signifikanz der Variation der Regressionskoeffizienten kann empirisch untersucht werden; das entsprechende Vorgehen wird weiter unten anhand eines Beispiels demonstriert.

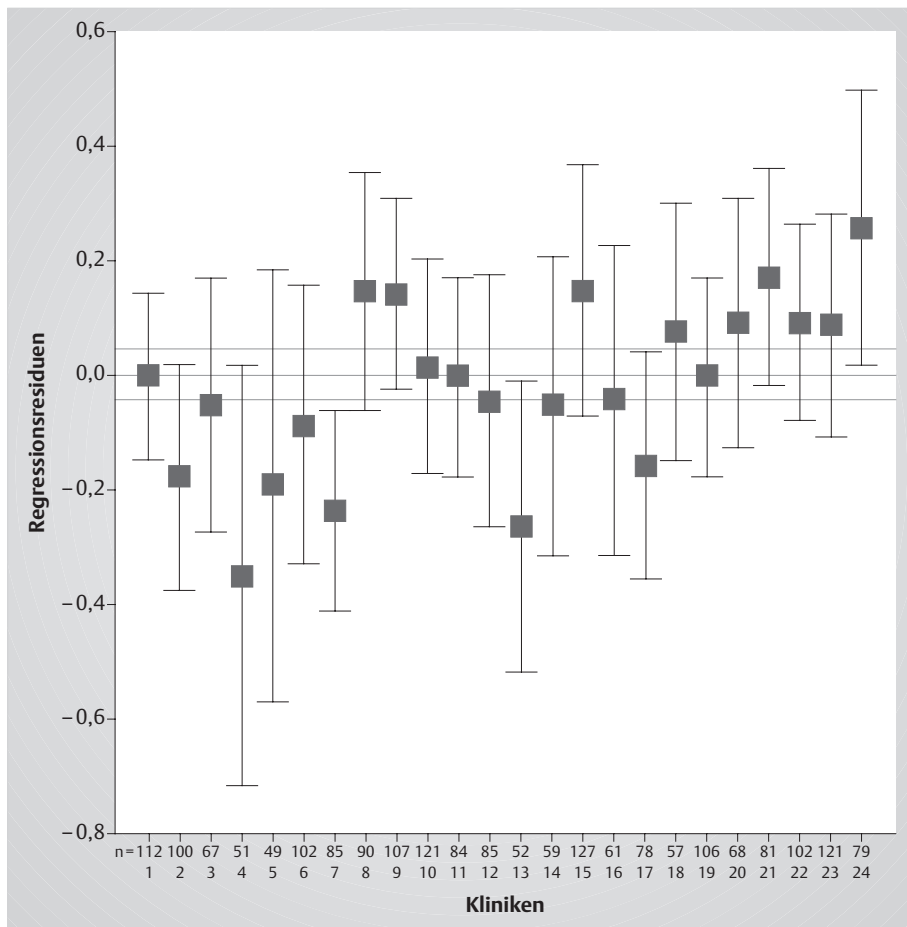


Abb. 2 Klinikvergleich auf der Basis von Regressionsresiduen (Fehlerbalkendiagramm mit 95%-Konfidenzintervallen, die waagerechten Linien in der Mitte geben das 95%-Konfidenzintervall des Gesamtmittelwerts wieder).

## 2. Die Benchmarkingfragestellung

Für Qualitätssicherungsprogramme ist die Benchmarkingfragestellung die zentrale: Wie unterscheiden sich Rehabilitationseinrichtungen generell über alle Patienten hinweg bei Kontrolle der relevanten Confounder? Um über einen kontrollierten, vom Einfluss der Confounder befreiten Wert zu verfügen, der grafisch visualisiert werden kann, werden häufig (vgl. z. B. [10–12]) die Residuen einer Regressionsanalyse mit Hilfe von Fehlerbalkendiagrammen klinikbezogen dargestellt. Abb. 2 veranschaulicht dies beispielhaft: Jede Einrichtung wird durch einen Fehlerbalken (Mittelwert mit 95%-Konfidenzintervall) dargestellt. Überlappt sich der Fehlerbalken einer Einrichtung nicht mit dem Fehlerbalken einer anderen Einrichtung bzw. nicht mit dem Konfidenzintervall des Gesamtmittelwerts, so bestehen signifikante Unterschiede zu der anderen Einrichtung bzw. zum Gesamtmittelwert.

Um über einen Residualwert zu verfügen, der über alle Einrichtungen hinweg vergleichbar ist und grafisch dargestellt werden kann, kann man nicht das oben erwähnte Random-Coefficient Regression Model anwenden, da unter diesem Modell jede Einrichtung eine spezifische Regressionsgleichung erhält, sodass die Residualwerte nicht zwischen Einrichtungen verglichen werden könnten. Man würde dadurch nach qualitätsbezogenen Differenzen (z. B. unterschiedliche Lagen des Intercepts, die unterschiedliche Niveaus von Reha-Effekten abbilden) adjustieren, obwohl diese Differenzen durch die grafische Darstellung der Residuen doch erst visualisiert werden sollen.

Erforderlich ist also die Annahme, dass die Regressionskoeffizienten über die Einrichtungen hinweg konstant sind, was sich empirisch mithilfe Hierarchischer Linearer Modelle prüfen lässt (vgl. das Beispiel weiter unten). Die Aufnahme relevanter Ebene-2-Prädiktoren empfiehlt sich, um möglichst die gesamte, nicht durch Qualitätsunterschiede der Einrichtungen bedingte Varianz zu erklären. Bei der Auswahl der Ebene-2-Prädiktoren gilt das eingangs bereits Gesagte, dass nur solche Variablen Berücksichtigung finden dürfen, die von der Klinik nicht direkt beeinflussbar sind und insofern keine Qualitätsaspekte des Leistungsgeschehens abbilden (z. B. mittleres Alter der Patienten einer Klinik, Anteil der männlichen Patienten, mittlere psychische Eingangsbelastung der Patienten). Für Benchmarkingfragestellungen würde man also ein HLM verwenden, welches sowohl Ebene-1- als auch Ebene-2-Prädiktoren berücksichtigt, aber keine Variabilität der Regressionskoeffizienten beinhaltet.

## 3. Die Zuweisungfragestellung

Bei der Frage, in welcher Klinik für einen bestimmten Patienten bzw. für eine bestimmte Merkmalskonstellation (z. B. 70-jährige Patientin mit hoher psychischer Belastung, aber ansonsten geringer Komorbidität) der beste Rehabilitationserfolg zu erwarten ist, ist die Berücksichtigung möglicher Unterschiede in der Bedeutung von Regressionskoeffizienten zwischen Einrichtungen erwünscht. Hier würde also der „Intercepts- and Slopes-as-Outcomes“-Ansatz, wie er z. B. in Gl. 4a–4c abgebildet wird, angewandt werden. Es soll für jede Einrichtung die optimale Prognose des Reha-Outcomes erfolgen. Dazu würden die obigen Merkmalswerte des zuzuweisenden Patienten in die einrichtungsspe-



zifischen Regressionsgleichungen eingesetzt und untersucht, bei welcher Einrichtung der prognostizierte Outcomewert optimal ist. Zu berücksichtigen sind dabei allerdings die insbesondere bei niedrigen Fallzahlen bedeutsamen Zufallsschwankungen der Regressionskoeffizienten und damit der Prognosewert sowie die Möglichkeit, dass sich im zeitlichen Verlauf Prädiktor-Kriteriums-Zusammenhänge in einer Einrichtung verändern können. Prognosemodelle sollten deshalb vor dem Einsatz in der Praxis einer Kreuzvalidierung unterzogen werden und in der Folgezeit regelmäßig auf Aktualisierungsbedarf geprüft werden.

### Ein Anwendungsbeispiel

Abschließend soll die Anwendung und Interpretation der Hierarchischen Linearen Modelle anhand von Daten aus dem Qualitätssicherungsprogramm der gesetzlichen Krankenkassen [1] verdeutlicht werden. In 32 orthopädischen Rehabilitationskliniken wurden mittels des IRES-Fragebogens [13] zu Aufnahme ( $t_0$ ) und Entlassung ( $t_1$ ) an ca. 150–200 Patienten pro Einrichtung der somatische (St0, St1), funktionale (Ft0, Ft1) und psychosoziale Status (Pt0, Pt1) erfasst. Ferner wurden Alter, Geschlecht (Sex) und – über einen ärztlicherseits zu bearbeitenden Dokumentationsbogen – Komorbidität (Komo) und Reha-Motivation (Motiv) erhoben. Im Sinne der oben skizzierten Prädiktorfragestellung sollen Einflussfaktoren des Reha-Outcomes im funktionalen Bereich erfasst werden. Bezüglich aller Regressionskoeffizienten wird eine Variation über die Einrichtungen zugelassen; als Prädiktoren auf der Ebene 2 werden das mittlere Alter der Patienten einer Klinik (AlterM), der mittlere Eingangsstatus auf der somatischen (SM), funktionalen (FM) und psychosozialen Ebene (PM) sowie der Anteil von männlichen Patienten (SEXM) betrachtet. Die Variablen zum somatischen, funktionalen und psychosozialen Status sind auf Patienten- und Klinikebene so gepolt, dass hohe Werte einen positiven Gesundheitsstatus und niedrige Werte eine Einschränkung der Gesundheit bedeuten.

Das Modell auf der Ebene 1 lautet damit:

$$\begin{aligned} Ft1_{ij} &= \beta_{0j} + \beta_{1j} \times \text{Alter}_{ij} + \beta_{2j} \times \text{Sex}_{ij} + \beta_{3j} \times \text{StO}_{ij} + \beta_{4j} \times \\ &\text{FtO}_{ij} + \beta_{5j} \times \text{PtO}_{ij} + \beta_{6j} \times \text{Komo}_{ij} + \beta_{7j} \times \text{Motiv}_{ij} + \Gamma_{ij} \end{aligned} \quad (\text{Gl. 6})$$

Die Gleichung 6 stellt – wie Gleichung 1 – das übliche Modell einer linearen Regression dar, nur dass nun insgesamt 7 Prädiktoren berücksichtigt werden (Alter, Geschlecht, somatischer, funktionaler und psychosozialer Eingangsstatus, Komorbidität und Reha-Motivation).

Auf der Ebene 2 lautet das Modell:

$$\beta_{qj} = \gamma_{q0} + \gamma_{q1} \times \text{AlterM}_j + \gamma_{q2} \times \text{SEXM}_j + \gamma_{q3} \times \text{SM}_j + \gamma_{q4} \times \text{FM}_j + \gamma_{q5} \times \text{PM}_j + u_{qj} \quad (\text{Gl. 7})$$

für  $q=0$  bis  $q=7$  (Intercept und sieben Steigungskoeffizienten, entsprechend der sieben in Gleichung 6 berücksichtigten Prädiktoren). Das Modell besteht also insgesamt aus acht Gleichungen, die jeweils zum Ausdruck bringen, dass die Größe der Koeffizienten abhängt von den Merkmalsausprägungen der Einrichtungen auf der Ebene 2 (mittleres Alter der Patienten, somatischer Status zu Reha-Beginn etc.). Der Wert  $u_{qj}$  stellt die klinikspezifische

Tab. 1 Ergebnisse des Hierarchischen Linearen Modells (berechnet mit HLM 5.05)

fester Effekt	Koeffizient	t-Wert	p-Wert
für das Intercept auf Ebene 1 ( $\beta_0$ )			
Intercept $\gamma_{00}$	-7,96	-3,74	0,001
SM $\gamma_{01}$	-0,55	-5,62	0,000
PM $\gamma_{02}$	0,35	2,67	0,013
ALTERM $\gamma_{03}$	0,15	4,72	0,000
für den Regressionskoeffizienten des Alters ( $\beta_1$ )			
Intercept $\gamma_{10}$	-0,01	-5,02	0,000
für den Regressionskoeffizienten des Geschlechts ( $\beta_2$ )			
Intercept $\gamma_{20}$	2,64	2,79	0,009
ALTERM $\gamma_{21}$	-0,04	-2,83	0,009
für den Regressionskoeffizienten des somatischen Status bei Aufnahme ( $\beta_3$ )			
Intercept $\gamma_{30}$	0,09	5,60	0,000
für den Regressionskoeffizienten des funktionalen Status bei Aufnahme ( $\beta_4$ )			
Intercept $\gamma_{40}$	1,75	6,33	0,000
ALTERM $\gamma_{41}$	-0,01	-3,66	0,001
für den Regressionskoeffizienten des psychosozialen Status bei Aufnahme ( $\beta_5$ )			
Intercept $\gamma_{50}$	0,09	5,11	0,000
für den Regressionskoeffizienten der Reha-Motivation ( $\beta_6$ )			
Intercept $\gamma_{60}$	-0,06	-3,14	0,004
für den Regressionskoeffizienten der Komorbidität ( $\beta_{60}$ )			
Intercept $\gamma_{70}$	-0,03	-4,09	0,000
Zufallseffekt	Varianzkomponente	p-Wert	
Intercept $u_0$	0,07242	> 0,500	
Alter $u_1$	0,00001	> 0,500	
Geschlecht $u_2$	0,03440	0,073	
somatischer Status $t_0$ $u_3$	0,00285	0,382	
funktionaler Status $t_0$ $u_4$	0,00520	0,008	
psychosozialer Status $t_0$ $u_5$	0,00336	0,087	
Reha-Motivation $u_6$	0,00128	> 0,500	
Komorbidität $u_7$	0,00039	> 0,500	

*kursiv* = Ebene-2-Prädiktoren; die Werte in den Spalten Koeffizient und t-Wert wurden auf zwei Stellen nach dem Komma gerundet

Abweichung des Koeffizienten  $\beta_q$  vom Mittelwert aller Einrichtungen dar. Der Wert  $\gamma_{q0}$  stellt den Wert des jeweiligen Koeffizienten  $\beta_q$  dar, der für eine Einrichtung zu erwarten ist, die auf allen berücksichtigten Ebene-2-Prädiktoren die Ausprägung Null aufweist und keine Abweichung vom Mittelwert aller Kliniken aufweist. Für eine bessere Interpretierbarkeit des  $\gamma_{q0}$  ist es aber auch möglich, eine Zentrierung der Prädiktoren am jeweiligen Gesamtmittelwert durchzuführen. Dann steht  $\gamma_{q0}$  für den Wert, der zu erwarten ist, wenn eine Einrichtung bezüglich aller Ebene-2-Prädiktoren (z. B. mittleres Alter der Patienten) den Mittelwert aller Kliniken aufweist.

Das obige Modell, welches einem „Intercepts-and-Slopes-as-Outcomes“-Ansatz entspricht, wurde mit der Software HLM 5.05 berechnet. Nachdem diejenigen Ebene-2-Prädiktoren, die nicht signifikant werden, aus dem Modell herausgenommen wurden, sieht das Resultat so wie in Tab. 1 dargestellt aus.

### Die zufälligen Effekte

Der untere Teil („Random Effect“) zeigt die Signifikanz der Zufallseffekte an, also der Variation der Regressionskoeffizienten über die Einrichtungen hinweg. Es zeigt sich, dass lediglich der Prädiktor „Funktionaler Status bei Aufnahme“ eine signifikante Variation ( $p < 0,05$ ) aufweist, bei allen anderen Regressionskoeffizienten ist der Zufallseffekt nicht bedeutsam von Null verschieden. Die Variation des Regressionskoeffizienten des Prädiktors funktionaler Status bei Aufnahme ist in diesem Beispiel Ausdruck der Tatsache, dass die Reha-Effekte auf der funktionalen Ebene in den Kliniken unterschiedlich ausfallen (in Einrichtungen mit geringen Effekten ist der Eingangswert ein sehr guter Prädiktor des Entlasswerts, in Kliniken mit hohen Effekten ist die Vorhersagekraft des Eingangswerts geringer).

### Die festen Effekte

Der obere Teil der Tabelle („Fixed Effect“) stellt die Regressionskoeffizienten der Ebene-1- und Ebene-2-Prädiktoren dar. Die Ebene-2-Prädiktoren sind kursiv gedruckt. Bezüglich der Ebene-1-Prädiktoren zeigt sich, dass alle hier aufgenommenen Faktoren (Alter, Geschlecht, somatischer, funktionaler, psychosozialer Status sowie Komorbidität und Reha-Motivation) einen Einfluss auf den Entlassstatus im funktionalen Bereich besitzen (Koeffizienten  $\gamma_{10}$  bis  $\gamma_{70}$ ). Der Einfluss der Ebene-2-Prädiktoren soll im Folgenden näher betrachtet werden: Das Intercept der Regressionsgleichung (d. h. das Niveau des vorhergesagten funktionalen Status bei Entlassung) ist umso höher (d. h. der Reha-Erfolg umso deutlicher), je geringer der mittlere somatische Eingangstatus der Patienten der Klinik ist, je höher der mittlere psychosoziale Eingangstatus ist und je höher das mittlere Alter der Patienten der Klinik ist (Signifikanz der Koeffizienten  $\gamma_{01}$ ,  $\gamma_{02}$  und  $\gamma_{03}$  in Tab. 1). Vereinfacht ausgedrückt bestehen also für die Behandlung eines Patienten auf der funktionalen Ebene günstige Kontextbedingungen, wenn die Mitpatienten psychisch stabil sind, aber eher älter und somatisch stark belastet sind. Man kann die Hypothese aufstellen, dass sich hier ein Spezialisierungseffekt abbildet, dergestalt, dass Kliniken, die viel Erfahrung mit älteren, auf der somatischen Ebene stark eingeschränkten Patienten haben, Therapiekonzepte realisieren, die auf der funktionalen Ebene besonders effektiv sind, während Kliniken, die psychisch belastete Patienten haben, eher Probleme haben, auf der funktionalen Ebene hohe Reha-Effekte zu erzielen – unter Umständen, weil sie therapeutische Kapazitäten verstärkt in die Behandlung der psychischen Komorbidität investieren.

Es ist zu beachten, dass das Alter des Patienten auf den Ebenen 1 und 2 einen entgegengesetzten Effekt besitzt: Das individuelle Alter des Patienten stellt einen „Risikofaktor“ dar (bei höherem Alter werden geringere Reha-Effekte erzielt), das mittlere Alter der Patienten einer Klinik hingegen einen „Protektivfaktor“ (bei höherem mittleren Alter werden höhere Reha-Effekte erzielt). Es ist also z. B. zu erwarten, dass ein 60-jähriger Patient in einer Klinik mit einem mittleren Patientenalter von 70 Jahren im funktionalen Bereich größere Fortschritte macht als in einer Klinik mit einem mittleren Patientenalter von 60 Jahren.

Das Alter als Ebene-2-Prädiktor ist auch im Hinblick auf zwei weitere Ebene-1-Prädiktoren von Bedeutung: Der Regressionskoeffizient des funktionalen Eingangstatus ( $\gamma_{40} + \gamma_{41} \times \text{ALTERM}$ , vgl. Tab. 1) ist für alle in den Kliniken vorkommenden Patienten-

altersmittelwerte (sie schwanken im vorliegenden Datensatz zwischen 61 und 72 Jahren) positiv; bei hohem mittleren Alter ist der Regressionskoeffizient jedoch geringer, d. h. in hohem Alter bestimmt der Eingangsstatus in geringerem Maße den Entlassstatus. Ferner besitzt das mittlere Patientenalter einen Einfluss auf den Regressionskoeffizienten des Geschlechts (Koeffizient  $\gamma_{21}$ ). Hier ist die Beziehung komplexer, weil das Vorzeichen des Regressionskoeffizienten je nach mittlerem Patientenalter wechselt. Um zu berechnen, bis zu welchem mittleren Patientenalter Frauen stärker von der Rehabilitation profitieren, wird die Berechnungsformel für den Regressionskoeffizienten des Geschlechts mit Null gleichgesetzt:

$$2,64 - 0,04 \times \text{AlterM} = 0 \quad (\text{Gl. 8})$$

$$\text{AlterM} = 66 \quad (\text{Gl. 9})$$

Das heißt, in Kliniken mit jüngeren Patienten (mittleres Patientenalter bis 66 Jahre) ist das Vorzeichen des Regressionskoeffizienten des Geschlechts positiv, was angesichts der Polung der Variable Geschlecht bedeutet, dass Frauen stärker von der Rehabilitation profitieren, während sich ab einem mittleren Patientenalter von 66 Jahren (das ist bei 84% der Einrichtungen der Fall) die Beziehung umkehrt und bei Frauen nach Kontrolle aller anderen Prädiktoren geringere Reha-Effekte erzielt werden.

Mit dem Beispiel sollte demonstriert werden, dass die mit hierarchischen Modellen mögliche Berücksichtigung von Prädiktoren unterschiedlicher Ebenen zu neuen Erkenntnissen führen kann. Ein Risikoadjustierungsmodell zum Einrichtungsvergleich auf der Basis des oben spezifizierten hierarchischen Modells kontrolliert mehr relevante Confounder und ist insofern elaborierter als ein regressionsanalytisches Verfahren, welches lediglich Ebene-1-Prädiktoren berücksichtigt.

Um die oben dargestellten Ergebnisse für konkrete Einrichtungsvergleiche nutzbar zu machen, würde man ein HLM berechnen, das sowohl Ebene-1- als auch Ebene-2-Prädiktoren enthält, aber keine Variation der Regressionskoeffizienten beinhaltet. Die aus diesem Modell resultierenden Residualwerte der Einrichtungen könnten dann – wie in Abb. 2 dargestellt – mit Hilfe von Fehlerbalkendiagrammen visualisiert werden.

### Diskussion und Zusammenfassung

Mit Hierarchischen Linearen Modellen lassen sich gleichzeitig Prädiktoren auf der Ebene der Patienten und auf der Ebene der Einrichtungen berücksichtigen. Dadurch können Risikoadjustierungen noch elaborierter durchgeführt werden. Vergleiche zwischen der Anwendung einfacher linearer Regressionsanalysen und Hierarchischer Linearer Modelle auf Daten aus Qualitätssicherungsprogrammen zeigen, dass die Anwendung der HLMs gegenüber dem üblichen Vorgehen auf der Basis einfacher linearer Regressionsanalysen nicht zu stark divergierenden Resultaten führt, sie kann jedoch nützlich sein, um die Vergleiche in Bezug auf Einrichtungen mit extremen Merkmalen in der Patientenstruktur „fairer“ zu gestalten und sollte insofern auch in externen Qualitätssicherungsprogrammen Berücksichtigung finden.

Bei der Anwendung von HLMs sind – ebenso wie bei einfachen Regressionsanalysen – gewisse Datenvoraussetzungen zu beachten, über die im Detail Raudenbush u. Bryk ([4], Kap. 9) berichten. Im Wesentlichen sind dies: keine zu hohe Multikollinearität, homogene Fehlervarianzen über die Einrichtungen hinweg und Normalverteilung der Residualwerte. Neben intervallskalierten Variablen können in ein HLM auch dummykodierte Variablen aufgenommen werden. Die Teststärke der Prüfung der Ebene-2-Prädiktoren hängt entscheidend von der Anzahl der Einrichtungen ab, die in die Analyse eingehen. Simulationsstudien (vgl. [5], S. 124 ff.) zeigen, dass bei 30 Einrichtungen mit einer Fallzahl von  $n = 30$  pro Einheit eine hinreichende Teststärke erzielt wird. Da in den Qualitätssicherungsprogrammen in der Rehabilitation pro Klinik meist deutlich mehr Fälle vorliegen (im obigen Datenbeispiel 150–200 Patienten pro Klinik), scheint in diesem Kontext eine Anwendung ab ca. 15 Kliniken vertretbar.

Schließlich ist zu berücksichtigen, dass die Anwendung der HLMs nur dann zu aussagekräftigen Ergebnissen der Risikoadjustierung führen kann, wenn alle wesentlichen Confounder des Rehabilitationserfolgs reliabel und valide erfasst wurden. Voraussetzungen für faire Einrichtungsvergleiche sind neben einer elaborierten Vergleichsmethodik auch eine theoriegeleitete bzw. auf empirischen Studien zu den Prädiktoren der Reha-Effekte basierende Auswahl der Confounder sowie der Einsatz methodisch abgesicherter Instrumente.

## Literatur

- <sup>1</sup> Farin E, Gerdes N, Jäckel WH, Follert P, Klein K, Glattacker M. „Qualitätsprofile“ von Rehabilitationskliniken als Modell der Qualitätsmessung in Einrichtungen des Gesundheitswesens. *Gesundheitsökonomie und Qualitätsmanagement* 2003; 8 (3): 191–204
- <sup>2</sup> Iezzoni LI. Risk adjusting rehabilitation outcomes. An overview of methodologic issues. *American Journal of Physical Medicine and Rehabilitation* 2004; 83: 316–326
- <sup>3</sup> Iezzoni LI. Risk adjustment for Measuring Healthcare Outcomes (3<sup>rd</sup> ed). Chicago: Health Administration Press, 2003
- <sup>4</sup> Raudenbush SW, Bryk AS. Hierarchical linear models. Applications and data analysis methods (2<sup>nd</sup> ed). Thousand Oaks, CA: Sage, 2002
- <sup>5</sup> Kreft I, Leeuw J de. Introducing Multilevel Modeling. London: Sage, 1998
- <sup>6</sup> Singer JD. Using SAS PROC MIXED to fit multilevel models, hierarchical models, and individual growth models. *Journal of Educational and Behavioral Statistics* 1998; 24 (4): 323–355
- <sup>7</sup> Raudenbush SW, Bryk A, Congdon R. HLM Hierarchical Linear and Nonlinear Modeling 2001 (Version 5.05).
- <sup>8</sup> Centre for Multilevel Modelling. MLwiN (Version 1.10.0007)
- <sup>9</sup> Wegscheider K. Methodische Anforderungen an Einrichtungsvergleiche. *Zeitschrift für Ärztliche Fortbildung und Qualitätssicherung* 2004; 98: 647–654
- <sup>10</sup> Schneeweiss S, Sangha O. Leistungsvergleiche in der Medizin. *Deutsche Medizinische Wochenschrift* 2001; 126: 918–924
- <sup>11</sup> Farin E, Glattacker M, Follert P, Kuhl C, Klein K, Jäckel WH. Einrichtungsvergleiche in der medizinischen Rehabilitation. *Zeitschrift für Ärztliche Fortbildung und Qualitätssicherung* 2004; 98: 655–662
- <sup>12</sup> Goldstein H, Spiegelhalter DJ. League tables and their limitations: Statistical issues in comparisons of institutional performance. *Journal of the Royal Statistical Society* 1996; 159 (3): 385–443
- <sup>13</sup> Bühlren B, Gerdes N, Jäckel WH. IRES-Indikatoren des Reha-Status. In: Schumacher J, Klaiberg A, Brähler E (Hrsg): *Diagnostische Verfahren zu Lebensqualität und Wohlbefinden*. Göttingen: Hogrefe, 2003: 180–183