

Auswirkungen der Fehlspezifikation von Meßfehlerkorrelationen auf die Parameterschätzung in linearen Strukturgleichungsmodellen in der Glaukomdiagnose

Orawa H, Martus P

Institut für Biometrie und Klinische Epidemiologie, Charité Universitätsmedizin Berlin, Berlin, Deutschland
helmut.orawa@charite.de

Einleitung Chronische Glaukome (Grüner Star) sind Erkrankungen, bei denen typische Schädigungen des Sehnervs auftreten, die Gesichtsfeldeinschränkungen zur Folge haben und ohne Behandlung zur Erblindung führen können. Glaukome sind keineswegs zwingend mit erhöhtem Augeninnendruck verbunden. Frühzeitige Diagnose ist beim Glaukom wichtig, da die Erkrankung vom Patienten erst in einer relativ fortgeschrittenen Phase wahrgenommen wird, aber einmal entstandene Schädigungen irreversibel sind.

Mangels eines *allgemeinen* Goldstandards bestehen mehrere quantitative diagnostische Referenzverfahren und eine Vielzahl neuerer Diagnosemethoden. Die Bewertung eines Diagnoseverfahrens geschieht herkömmlicherweise, indem es mit i. a. mehreren Referenzverfahren verglichen wird, die unterschiedliche Krankheitsaspekte bzw. Stadien des Krankheitsprozesses abbilden [1]. Dieses Vorgehen produziert voneinander abweichende Bewertungen. Wird ein neues Verfahren einer Referenz gegenübergestellt, die den gleichen Krankheitsaspekt mißt, führt dies zur Überbewertung eines Diagnoseinstruments. Der gleiche Effekt tritt verstärkt auf, wenn eine nicht beobachtete Störgröße, die nichts mit der Erkrankung zu tun hat, beide Verfahren beeinflusst (z.B. Konzentrationsfähigkeit bei psychophysischer Testung). Messen beide Verfahren unterschiedliche Krankheitsaspekte, resultiert dagegen u. U. eine relative Unterbewertung eines neuen Verfahrens.

Material und Methoden Es wurden Daten des Erlanger Glaukomregisters ausgewertet, das als klinisches Register zur Langzeitbeobachtung von Patienten mit manifestem Glaukom und Verdacht auf Glaukom angelegt ist. 237 Patienten und 140 nichterkrankte Kontrollen durchliefen vollständig eine Reihe von Diagnoseverfahren, die einer der drei Gruppen zuordenbar sind: Morphologische, psychophysische und elektrophysiologische Verfahren. Diese sind durch unterschiedliche Formen der Reizgebung und Reaktionserfassung charakterisiert und weisen zudem einen unterschiedlichen Grad von Objektivität auf. Die klinische Diagnose zur Klassifikation in die Gruppe der Erkrankten bzw. Nichterkrankten wurde in diesem Fall mittels der qualitativen Untersuchung des Sehnervkopfes durch zwei erfahrene Ophthalmologen gestellt.

Hier werden Lineare Strukturgleichungsmodelle (SEMs) verwendet, die insbesondere in der medizinischen Diagnostik gut anwendbar sind [2]. Denn bei der Diagnose von Krankheiten handelt es sich um die Bestimmung von Phänomenen, die sich häufig unmittelbarer Messung (Beobachtung) entziehen und vielfach nur mittelbar über verschiedene Indikatoren (diagnostische Meßverfahren) feststellbar sind. Strukturgleichungsmodelle bestehen aus einem Pfadmodell von i. a. nicht meßbaren, also latenten Variablen sowie zugeordneten Meßmodellen. Bei diesen Meßmodellen wird jede manifeste Variable genau einer latenten Variable zugeordnet. SEMs ermöglichen die Bewertung von Referenz- und neuen Diagnoseverfahren unter Berücksichtigung von Störgrößen und zeigen auf, wie gut diese den globalen Glaukomschaden quantifizieren. Ein individueller Patientenindex für den Glaukomschweregrad auf Basis von SEMs scheint die beste Annäherung an einen allgemeinen Goldstandard zu sein. Voraussetzung ist, daß Meßverfahren aus *unterschiedlichen* Gruppen von Verfahren (s. o.) zur Verfügung stehen.

Das Problem der Über- oder Unterbewertung von Diagnoseverfahren und damit auch eine fehlerhafte Quantifizierung des individuellen Glaukomschweregrades kann auch in SEMs auftreten. Es tritt als Folge der falschen Annahme auf, daß bestimmte Meßverfahren - bedingt auf den Glaukomschweregrad - unabhängig, oder m. a. W., daß die Meßfehler von Diagnoseverfahren unkorreliert seien. Zur Klärung der Frage, ob diese Annahme zutreffend ist, stehen mehrere Ansätze zur Verfügung. Eine Option ist der Vergleich der Korrelationsstruktur der Meßfehler in der erkrankten und nichterkrankten Teilstichprobe. Die Richtigkeit einer Modellannahme über die Korrelationsstruktur von Meßfehlern ist jedoch schwierig zu überprüfen. Daher wird versucht, die möglichen verzerrenden Auswirkungen von fehlerhaften Annahmen (Fehlspezifikation) auf die Bewertung der Meßverfahren (und damit letztlich auf die Einschätzung des individuellen Schweregrades) in ihrer Richtung und Größenordnung abzusichern. ~~Abzugeschrieben~~ dieses Projekts ist ein publiziertes Modell der konfirmatorischen Faktoranalyse (CFA), die einen Spezialfall eines SEM darstellt [3]. Es wird hier nicht in Gleichungsform sondern grafisch wiedergegeben (siehe Abb.).

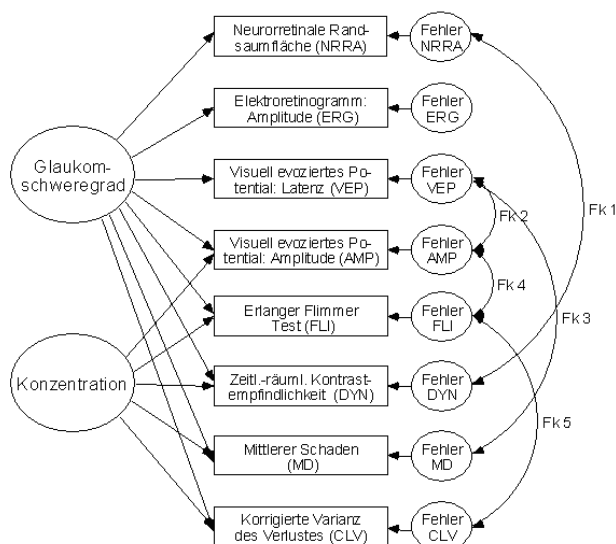


Abb.: Zwei-Faktoren-Modell für Diagnoseverfahren mit abhängigen Meßfehlern

Veranschaulicht werden 8 Diagnoseverfahren (Rechtecke), die - von oben nach unten - den o. g. Gruppen wie folgt zugeordnet sind: Verfahren 1: morphologische, Verfahren 2-4: elektrophysiologische und Verfahren 5-8: psychophysische. Die Diagnoseinstrumente werden einerseits von zwei nichtbeobachtbaren Faktoren (Ellipsen) beeinflusst, dem Glaukomschweregrad und der Konzentrationsfähigkeit der Patienten bei der Absolvierung diagnostischer Tests. Andererseits beeinflussen Meßfehlerkomponenten (Kreise) die Instrumente. Die Einflüsse werden durch einfache Pfeile (Pfade) dargestellt und angenommene Korrelationen zwischen den Meßfehlerkomponenten durch Doppelpfeile symbolisiert. In diesem Modell werden fünf Fehlerkorrelationen (Fk) angenommen, so daß - bis auf ERG - von abhängigen Meßverfahren ausgegangen wird. Für das Modell wurden Pfadkoeffizienten geschätzt, die das Maß der Beeinflussung angeben.

Bislang wurden auf Basis des abgebildeten Modells Bootstrapanalysen zur Schätzung der Standardfehler der Pfadkoeffizienten mit der Software AMOS durchgeführt. In der Analyse mit Amos kann nicht nach der Abhängigkeit adjustiert werden, die zwischen den beiden Augen eines Patienten besteht. Bei der Interpretation des globalen Tests für die Güte der Anpassung des Modells an die Daten wurde diese Abhängigkeit daher bislang in Form eines Korrekturfaktors berücksichtigt.

Ergebnisse Die Abschätzung des durch Modellfehlspezifikation induzierten Bias der interessierenden Pfadkoeffizienten wird in einem dreistufigen Projekt durchgeführt. Die Analyse der aufgrund von Modellvorgaben erzeugten deterministischen Daten (A) in AMOS bildet die Grundlage der Untersuchung von mit Hilfe zufälliger Störgrößen erzeugten Simulationsdaten (B) in SAS. Die Ergebnisse der Simulation werden in der Analyse der o. g. Registerdaten (C) angewendet: Mittels eines Bootstrapverfahrens zur Schätzung der Standardfehler der Pfadkoeffizienten wird die Robustheit des Modells der CFA überprüft. Die Analyse in SAS erlaubt hierbei eine angemessene Berücksichtigung der paarigen Natur der Daten.

Diskussion Das in B (Simulation) erstellte Instrumentarium erweist sich für die Überprüfung der Robustheit in C als brauchbar. Die Weiterentwicklung des CFA-Modells hinsichtlich der Adjustierung nach der Paarigkeit der Daten führt zu einer höheren diagnostischen Aussagekraft.

Literatur

- [1] Martus P. et al. Multivariate Approach for quantification of morphologic and functional damage in glaucoma. Invest Ophthalmol Vis Sci 2000; 41: 1099.
- [2] Bentler PM, Stein JA. Structural equation models in medical research. Stat Methods Med Res 1992; 1: 171.
- [3] Martus P. A measurement model of disease severity in absence of a gold standard. Methods Inf Med 2001; 40: 269.