

5 Material und Methoden

5.1 Die SMILE Studie

5.1.1 Aufbau

Die SMILE-Studie ist eine klinisch kontrollierte randomisierte doppelblinde Multicenter-Studie zur Überprüfung der Sicherheit und Wirksamkeit der Anwendung von Botulinum Toxin A (BoNT-A) zur Therapie von periokulären Lachfalten (Krähenfüßen).

Die SMILE-Studie wurde finanziert von der Ipsen Pharma GmbH (Ettlingen, Deutschland).

5.1.2 Entwicklung des Schulungsatlas

Bei einem Treffen im Januar 2004 in Paris wurden vier Experten standardisierte Aufnahmen von Patienten mit periokulären Falten unterschiedlicher Ausprägung vorgelegt. Aus diesen Fotos wählten die Experten im Konsensus-Verfahren für jeden Schweregrad der zu verwendenden Skalen exemplarisch jeweils drei Aufnahmen im entspannten und drei im angespannten Zustand aus. Die so ausgewählten und konsentierten Fotos, insgesamt 24, wurden mit entsprechender Einteilung zu einem Atlas zusammengefasst (siehe Anlage 1).

5.1.3 Studie zur Reproduzierbarkeit unter den Prüfern

Im Vorfeld der SMILE-Studie wurden im April 2004 neun Prüfer (vier Männer und fünf Frauen) mit Hilfe des zuvor erstellten Atlas in der Beurteilung von Lachfalten mit den o.g. Vier-Punkte-Skalen geschult. Im Anschluss an die Schulung wurden den Teilnehmern nacheinander insgesamt 98 standardisierte Fotografien von Lachfalten unterschiedlich starker Ausprägung zur Bewertung gezeigt. Dabei wurde zwischen Aufnahmen in entspanntem Zustand (50 Fotos) und in angespanntem Zustand (48 Fotos) unterschieden. Die Fotos wurden jeweils in zufälliger Reihenfolge gezeigt. Die Bewertungen der Teilnehmer wurden auf einem Bogen schriftlich erfasst. Die Teilnehmer durften für die Bewertung den Atlas zu Hilfe nehmen.

Von den ursprünglich vorgesehenen 98 Fotografien musste nachträglich eine Aufnahme in entspanntem Zustand herausgenommen werden, da die fotografierte Person Make-up verwendet hatte.

Am darauf folgenden Tag wurden dieselben Fotografien den Teilnehmern erneut zur Beurteilung vorgelegt.

5.1.4 Studie zur Reproduzierbarkeit unter den Experten

Nach Abschluss der SMILE-Studie wurden im Mai 2005 vier internationale Experten (drei Frauen und ein Mann) ebenfalls mit o.g. Atlas geschult. Anschließend wurden den Experten ebenfalls die insgesamt 97 standardisierten Fotografien von periokulären Lachfalten, 49 in entspanntem und 48 in angespanntem Zustand, zur Bewertung in zufälliger Reihenfolge vorgelegt. Für die Bewertung durften die Teilnehmer wiederum den Atlas zu Hilfe nehmen.

Zwei Tage später wurden den vier Teilnehmern dieselben Fotografien erneut zur Beurteilung gezeigt.

5.2 Statistische Maße zur Beurteilung der Reproduzierbarkeit kategorischer Variablen

5.2.1 Kappa-Koeffizient nach Cohen

Bei nur zwei Beobachtern, die mehrere Patienten anhand einer Skala beurteilen, kann man den Kappa-Koeffizient nach Cohen (Cohen 1960) verwenden. Als Basis dienen hierbei die konkordanten Urteile der beiden Beobachter. Zusätzlich zu den übereinstimmenden Urteilen werden außerdem die zufällig zu erwartenden Übereinstimmungen berücksichtigt:

$$\kappa = \frac{P_{obs} - P_{exp}}{1 - P_{exp}}.$$

Die zufällig zu erwartenden Übereinstimmungen sind der Anteil der übereinstimmenden Beurteilungen, der dem Anteil entspricht, der zu erwarten ist, wenn die Beurteilungen der beiden Beobachter (im statistischen Sinne) unabhängig sind.

Dieser Anteil wird berechnet nach:

$$p_{\text{exp}} = \frac{\sum_{i=0}^k e_{ii}}{N}.$$

Die e_{ii} werden aus den Randsummen folgendermaßen berechnet:

$$e_{ii} = \frac{f_i \cdot g_i}{N} \quad (i = 0, \dots, k).$$

Dabei sind f_i und g_i die Häufigkeiten, mit denen die beiden Beobachter den Skalen-Wert i vergeben haben.

p_{obs} bezeichnet den beobachteten Anteil der konkordanten Urteile.

p_{exp} bezeichnet den Anteil der konkordanten Urteile, die rein zufällig erwartet werden.

Kappa kann folgende Werte annehmen:

$\kappa = 1$, falls $p_{\text{obs}} = 1$ (d. h. alle Urteile stimmen überein).

$\kappa = 0$, falls $p_{\text{obs}} = p_{\text{exp}}$ (d. h. die Anzahl der konkordanten Urteile entspricht der Zufallserwartung).

Kappa kann theoretisch auch negative Werte annehmen. Dieser Fall tritt jedoch nur ein, wenn sich die beiden Beurteiler bemühen, absichtlich unterschiedliche Skalen-Werte zu vergeben und hat demzufolge praktisch keine Bedeutung.

Nach Altman (Altman 1998) gilt für alle Kappa-Koeffizienten, dass Werte zwischen 0,81 und 1,0 für eine sehr gute, Werte zwischen 0,61 und 0,8 für eine gute und Werte zwischen 0,41 und 0,6 für eine mäßige, Werte zwischen 0,21 und 0,40 für eine schlechte und Werte $< 0,2$ für eine sehr schlechte Übereinstimmung sprechen.

Der Nachteil dieses Kappa-Koeffizienten besteht darin, dass er ungewichtet ist. Das heißt, dass bei den diskordanten Urteilen der Grad der Diskordanz nicht berücksichtigt wird.

5.2.2 Gewichteter Kappa-Koeffizient nach Cohen

Der gewichtete Kappa-Koeffizient nach Cohen (Cohen 1968) berücksichtigt bei diskordanten Urteilen neben der Diskordanz an sich auch deren Grad. Er trägt der Bedeutung und den Folgen einer möglichen Fehlbeurteilung in der späteren Anwendung der Skalen Rechnung. Das Prinzip besteht darin, dass einer großen Diskrepanz der Urteile ein höheres Gewicht beigemessen wird und einer geringen Diskrepanz ein niedrigeres Gewicht. Die Basis bilden hierbei die diskordanten Urteile. Wenn die diskordanten Urteile überwiegend niedrige Gewichte haben und somit wenig gravierend sind, wirkt sich dies günstig auf κ aus. Der gewichtete Kappa-Koeffizient nach Cohen fällt dann etwas größer aus als der oben beschriebene ungewichtete Kappa-Koeffizient nach Cohen. Voraussetzung für diesen Koeffizienten ist allerdings, dass die zur Auswahl stehenden Kategorien mindestens ordinalskaliert sind.

Wie der ungewichtete Kappa-Koeffizient nach Cohen kann auch der gewichtete Kappa-Koeffizient nur bei zwei Beobachtern angewandt werden.

Die Diskrepanzgewichte sollten vor der Untersuchung, nach Möglichkeit durch einen Konsensusbeschluss von Experten, festgelegt werden oder es sollte auf standardisierte Gewichte zurückgegriffen werden. In dieser Arbeit wurden die von Cicchetti und Allison vorgeschlagenen standardisierten Gewichtungen verwendet (Cicchetti u. Allison 1971).

Dem gravierendsten Unterschied wird dabei üblicherweise das Gewicht 1 zugeordnet, konkordante Urteile erhalten das Gewicht 0. Alle anderen Diskrepanzen erhalten je nach ihrer Bedeutung Gewichte zwischen 0 und 1. Die Formel für den gewichteten Kappa-Koeffizienten lautet:

$$\kappa_w = 1 - \frac{\sum_{i,j=0}^k w_{ij} f_{ij}}{\sum_{i,j=0}^k w_{ij} e_{ij}} .$$

Die Häufigkeit f_{ij} ($i, j=0, \dots, k$) ist die Anzahl der Patienten, denen der Beobachter A die Ausprägung i und der Beobachter B die Ausprägung j zugewiesen hat. Die w_{ij} sind die Diskrepanzgewichte, die zwischen 0 und 1 liegen sollten. $w_{ij}=1$ bezeichnet demnach die größtmögliche Diskrepanz. Außerdem gilt sinnvollerweise $w_{ij}=w_{ji}$ und $w_{ij}=1=0$. In den gewichteten κ -

Koeffizienten fließen also nur die diskordanten Urteile ein. Die erwarteten Häufigkeiten berechnen sich nach:

$$e_{ij} = \frac{f_i \cdot g_j}{N}.$$

Dabei sind f_i und g_j die Häufigkeiten, mit denen die beiden Beobachter den Wert i bzw. j vergeben haben.

5.2.3 Kappa-Koeffizient nach Fleiss

Der Kappa-Koeffizient nach Fleiss (Fleiss 1971) kann für mehrere Beurteiler angewandt werden. Basis sind auch hier wie beim ungewichteten Koeffizienten von Cohen die konkordanten Urteile. Auch hier wird die Zufallserwartung mitbestimmt und in die Berechnung miteinbezogen. Der Kappa-Koeffizient nach Fleiss kann die gleichen Werte annehmen, die auch für den Kappa-Koeffizienten nach Cohen gelten.

5.3 Datenmanagement / Statistik

Die Datenverwaltung erfolgte mit einer Accessdatenbank, die Datenauswertung mit SAS[®] (Version 9.1).

Für die Auswertungen wurden folgende Maßzahlen berechnet:

- Zur Bestimmung der interindividuellen Variabilität: Kappa-Koeffizient nach Fleiss
- Zur Bestimmung der intraindividuellen Variabilität: Kappa-Koeffizient nach Cohen und der gewichtete Kappa-Koeffizient nach Cohen mit Gewichtungen nach Cicchetti-Allison

Für den gewichteten Kappa-Koeffizienten wurden die folgenden Wichtungsfaktoren verwendet: 0 für konkordante Urteile, 1/3 für die Diskordanzen 0-1, 1-2 und 2-3, 2/3 für die Diskordanzen 0-2 und 1-3 und 1 für die größtmögliche Diskordanz 0-3.