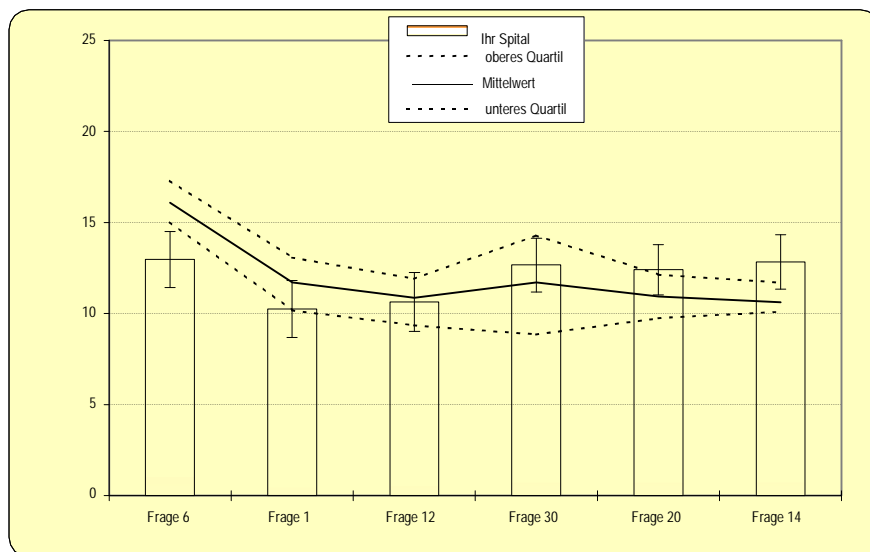


Bedeutung der Signifikanz von Unterschieden

Bevor intervenierende Maßnahmen eingeleitet werden, besteht die Möglichkeit zu klären, ob die im externen Vergleich gefundenen Unterschiede des eigenen Spitals zum Mittelwert aller Spitäler aussagekräftig sind oder ob diese Mittelwertunterschiede möglicherweise rein „zufällig“ zustande gekommen sind. Anhand der nachfolgenden Graphik soll diese Frage diskutiert werden:



Grafik : Unzufriedenheits-Index (0 = max. Zufriedenheit ; 100 = max. Unzufriedenheit)

A

	Frage 6	Frage 1	Frage 12	Frage 30	Frage 20	Frage 14
Wert für Ihr Spital	13.0	10.2	10.6	12.7	12.4	12.8
Mittelwert aller Spitäler in der Auswertung	16.1	11.7	10.9	11.7	10.9	10.6
oberes Quartil	17.3	13.1	11.9	14.3	12.1	11.7
unteres Quartil	15.0	10.2	9.3	8.8	9.7	10.1
Vertrauensintervall (90 %) + / -	1.5	1.6	1.6	1.5	1.4	1.5
Standardabweichung	20.7	21.1	21.7	19.3	18.5	20.1
Anzahl Angaben (für Ihr Spital)	490	493	484	464	480	491

B

Interpretation der Vertrauensintervalle:	VI deutlich unterhalb Mittelwert	VI überlappt ganz knapp mit MW	Spitalwert etwa gleich Mittelwert	VI überlappt deutlich mit Mittelwert	VI überlappt knapp nicht mit MW	VI deutlich überhalb Mittelwert
--	----------------------------------	--------------------------------	-----------------------------------	--------------------------------------	---------------------------------	---------------------------------

C

statist. Verfahren : t-Test	p = 0.00	p = 0.11	p = 0.73	p = 0.29	p = 0.08	p = 0.01
Interpretation: (Annahme: Signifikanzniveau p <= 0.05)	Unterschied klar signifikant	Unterschied knapp nicht signifikant	Unterschied klar nicht signifikant	Unterschied klar nicht signifikant	Unterschied knapp nicht signifikant	Unterschied klar signifikant

Darstellung 2 : Auswertung ausgewählter Fragen für das Musterspital (N=500) im externen Vergleich mit einem Datenpool aus 15 Spitälern (N=8056)

Teil A: Darstellung in der Detailauswertung

Teil B: Interpretation der Resultate der Detailauswertung

Teil C: statistische Untersuchung der Signifikanz der Unterschiede mittels t-Test

Begriffsdefinitionen

Im Folgenden werden der Signifikanztest, das Signifikanzniveau und der dazugehörige p-Wert näher erläutert und dem Vertrauensintervall (Konfidenzintervall) gegenübergestellt. Mit diesem Vergleich sollen Gründe aufgezeigt werden, für deskriptive Zwecke einer Studie, der Darstellung des Vertrauensintervalls gegenüber dem p-Wert den Vorzug zu geben.

Signifikanztest, Signifikanzniveau, p-Wert:

Mit dem Signifikanztest kann man a priori formulierte Hypothesen überprüfen. Die Nullhypothese (H_0) ist meistens so formuliert, dass kein Effekt postuliert wird (der Mittelwert der Zufriedenheit der Patienten aus Klinik A ist **gleich** wie die Zufriedenheit der Patienten aller anderen Kliniken). Die Alternativhypothese (H_1) formuliert einen Effekt bezüglich eines interessierenden Merkmals (der Mittelwert der Zufriedenheit der Patienten aus Klinik A **unterscheidet** sich von dem der Patienten aller anderen Kliniken). Ein häufig verwendeter Signifikanztest ist der t-Test. Das Ergebnis des t-Tests ist der p-Wert. Dann wird eine Irrtumswahrscheinlichkeit festgelegt (Signifikanzniveau α). Ist der p-Wert **kleiner** als das festgelegte Signifikanzniveau, so liegt statistische Signifikanz zum Niveau α vor, das heißt, die Nullhypothese wird abgelehnt. Die Wahrscheinlichkeit, dass die Nullhypothese fälschlicherweise abgelehnt wird, liegt je nach festgelegtem Signifikanzniveau bei 5% oder bei 1% [1]. Ein kleiner p-Wert sagt aus, dass es statistische Evidenz für einen Unterschied gibt. Wie groß dieser Unterschied ist, kann man am p-Wert nicht ablesen. Für diesen Zweck muss man die Größe des geschätzten Unterschiedes interpretieren, am besten im Zusammenhang mit einem Vertrauensintervall [1].

Das Vertrauensintervall (Konfidenzintervall)

Ein Vertrauensintervall (Konfidenzintervall) ist ein Unsicherheitsbereich für die Schätzung eines interessierenden Parameters (z.B. Mittelwert, Median, Differenz zweier Mittelwerte) aus einer Stichprobe. Ein 90% Vertrauensintervall beispielsweise enthält den gesuchten Parameter mit einer Wahrscheinlichkeit von 90% [2]. Aus einem Vertrauensintervall lassen sich Rückschlüsse auf die statistische Signifikanz ableiten. Um wie viel kleiner der p-Wert ist, kann man am Vertrauensintervall nicht ablesen. Dafür muss man den p-Wert selbst über die entsprechende Teststatistik berechnen. Die Angabe eines Vertrauensintervalls macht somit die Angabe eines p-Werts nicht völlig überflüssig [3]. Der p-Wert ist die genaue Angabe der statistischen Evidenz, das Vertrauensintervall liefert interpretierbare Informationen über die Unsicherheit des beobachteten Effekts.

Ein Vertrauensintervall liefert (im Gegensatz zum p-Wert) Informationen in der gleichen Skala wie die untersuchte Variable. Wird zum Beispiel die Zufriedenheit der Patienten untersucht, so ist die Einheit der Werte des Vertrauensintervalls ebenfalls ein Zufriedenheitswert [1].

¹ Bender R, Lange S. Was ist der P-Wert? Dtsch med Wschr 2001; 126: T39-T40.

² Bender R, Lange S. Was ist ein Konfidenzintervall? Dtsch med Wschr 2001; 126: T40.

³ Gardner MJ, Altman DG. Confidence intervals rather than P values: Estimating rather than hypothesis testing. Brit med J 1986; 292: 746-750.

Diskussion

Den obengenannten Vorzügen des Vertrauensintervalls für die deskriptive Statistik werden in der vorliegenden Detailauswertung Rechnung getragen.

Das Beispiel in der Darstellung 2 soll anschaulich zeigen, dass der Signifikanzwert und das Vertrauensintervall optisch miteinander in Beziehung stehen. Darstellung 2 lässt sich folgendermaßen interpretieren: Zum Mittelwert der einzelnen Fragen (bezüglich eines Spitals) wird ein Vertrauensintervall ausgegeben. Wenn das Vertrauensintervall des einzelnen Spitals den Mittelwert, der von allen Spitalern erzeugt wurde, berührt, also sehr nahe am Mittelwert des Gesamtpools liegt oder diesen überlappt, so weist der t-Test kein signifikantes Ergebnis aus. Umgekehrt wird ein signifikantes Ergebnis bei den Fragen erzielt, deren Vertrauensintervall den Mittelwert nicht berührt. Diese Beispiele zeigen klar, dass auch ohne Berechnung des p-Wertes, anhand des Vertrauensintervalls Aussagen über die Signifikanz eines Unterschiedes gemacht werden können.

Auf Wunsch können wir selbstverständlich den p-Wert für jedes Spital berechnen. Dadurch entsteht ein Mehraufwand, den wir verrechnen müssen.

Ein signifikantes Ergebnis bedeutet noch nicht zwingend ein relevantes Ergebnis. Nur eine Diskussion der Ergebnisse gibt Aufschluss darüber, durch welche Drittvariablen dieses Ergebnis allenfalls zusätzlich beeinflusst wird (nicht alle Einflussvariablen können in einer Erhebung mitberücksichtigt werden).

Für die „richtige“ Interpretation der Resultate ist es nötig:

- wichtige strukturelle und personenbezogene Besonderheiten (z.B. Stadt/Land-Gegensatz, Schweregrad der Krankheiten, Erwartungshaltung der Patienten) zu berücksichtigen.
- zu berücksichtigen, dass jede Messung eine statistisch beschreibbare Unschärfe aufweist. Die gelieferte Darstellung ist unserer Meinung nach genügend genau für eine gute Interpretation dieses Aspektes.