

Analyse von Sicherheitsklima-Befragungen: Andere Methode – anderes Resultat?!

69

Giai J, Boussat B, Occelli P, Gandon G, Seigneurin A, Michel P, François P et al.: Hospital survey on patient safety culture (HSOPS): variability of scoring strategies

International Journal for Quality in Health Care 2017; doi: 10.1093/intqhc/mzx086

Der Einschätzung des Sicherheitsklimas durch die Mitarbeitenden eines Spitals kommt eine wichtige Bedeutung zu. Typischerweise werden diese Einschätzungen über schriftliche Befragungen erhoben. Der *Hospital Survey on Patient Safety (HSOPS)* ist eines der international häufig eingesetzten Fragebogen-Instrumente. Der HSOPS beinhaltet 42 Fragen, die 12 verschiedenen Dimensionen zugeordnet sind, zum Beispiel „Teamwork in der Abteilung“ oder „Unterstützung durch die Spitalleitung“. Die Fragen werden auf einer 5-stufigen Skala beantwortet, die von „stimme gar nicht zu“ bis „stimme voll zu“ verläuft. Für die Aggregation der Antworten einer Gruppe (z.B. Mitarbeitende einer Abteilung) existieren verschiedene Methoden, z.B. der Durchschnitt der individuellen Mittelwerte. Basierend auf der Aggregation können dann Daten verglichen werden, z.B. der Vergleich von Dimensions-Mittelwerten zwischen Abteilungen. Giai et al. untersuchten die Frage, ob unterschiedliche Verfahren der Aggregation der Sicherheitsklima-Daten zu unterschiedlichen Ergebnissen führen. Diese Frage ist wichtig, da jeder Vergleich von aggregierten Klima-Daten davon betroffen sein kann. Wenn beispielsweise die Aussage, ob sich das Klima einer Abteilung aus Sicht der Mitarbeitenden verbessert oder verschlechtert hat, stark von der Auswertungsmethode abhängt, sind solche Vergleiche ohne Normierung der Methode zweifelhaft. Die Autoren verwendeten für ihre Untersuchung die Befragungsdaten der Mitarbeitenden eines grossen französischen Universitätsspitals mit 14 Abteilungen (n=3'978). Sie wendeten drei verschiedene Methoden an, um die Antworten der Mitarbeitenden auszuwerten: 1) Der Anteil der Mitarbeitenden, die eine positive Antwort gaben, gemittelt über die Fragen einer Dimension. Dieses Verfahren beruht auf der Dichotomisierung der 5-stufigen Antwort-Skala und gilt als leichter interpretierbar; 2) Der Durchschnitt der individuellen arithmetischen Mittelwerte über die Fragen einer Dimension; 3) Die Summe der individuellen Antworten zu den Fragen einer Dimension, gemittelt über die Individuen. Alle Ergebnisse wurden dann auf %-Punkte normalisiert, um die Werte direkt vergleichbar zu machen. Im zweiten Schritt produzierten die Autoren anhand der Resultate mit jeder dieser Methoden und für jede Dimension eine Rangliste der Abteilungen des Spitals. So ergeben sich beispielsweise 3 Ranglisten für die Dimension „Teamwork“ die jeweils die 14 Abteilungen auflisten. Theoretisch sollten alle Methoden das gleiche Ranking von Abteilungen ergeben. Die Autoren analysierten auch, ob die Unterschiede zwischen den Methoden innerhalb einer Abteilung von Bedeutung sind, z.B. wenn sich die Verhältnisse der Dimensionen zueinander verändern. So könnte beispiels-

weise eine Dimension gegenüber einer zweiten Dimension mit der einen Methode deutlich positiver ausfallen als mit einer anderen. Dies würde die Ableitung von Problemfeldern in Frage stellen. Die Autoren zeigen, dass die Bedeutung der Ergebnisse erheblich und in vielfältiger Hinsicht von der gewählten Aggregationsmethode abhängt. Methode 1 führt zu grundsätzlich höheren Bewertungen als die anderen Verfahren. So gibt es für die Dimension „Unterstützung durch die Spitalleitung“ für eine Abteilung mit Methode 1 einen Wert von 14.6%, mit Methode 3 liegt dieser Wert bei 47.5%. Die Unterschiede zwischen den Methoden fallen je nach Bewertungs-Niveau unterschiedlich stark aus. Es handelt sich also nicht um eine einfache lineare Verschiebung. Auch Unterschiede in den Bewertungen zwischen Berufsgruppen können mit der einen Methode deutlich ausfallen, während sie mit einer anderen Aggregation unauffällig sind. Die Ranglisten der Abteilungen unterschieden sich erheblich zwischen den Auswertungsmethoden. So erzielte eine Abteilung mit Methode 1 den Rang 4/14, mit Methode 2 den Platz 13/14. Besonders grosse Unterschiede ergeben sich zwischen Methode 1 und 2 oder 3, was auf die Reduktion der Varianz durch die Dichotomisierung in Methode 1 zurückzuführen ist. Zwischen den Methoden 1 und 2 gab es im Median 14 absolute Rangunterschiede. Diese Untersuchung macht deutlich, dass die Art der Auswertung von Sicherheitsklima-Daten einen erheblichen Einfluss auf die Resultate und die sich ergebende Interpretation hat. Besonders die Auswertung als Anteil positiver Antworten (Dichotomisierung) ist kritisch zu sehen. Zunehmend werden die Ergebnisse von Sicherheitsklima-Befragung für interne Vergleiche (z.B. zwischen Abteilungen oder im Zeitverlauf) aber auch für externe Vergleiche eingesetzt. Es ist vorstellbar, dass die Bewertung des Sicherheitsklimas durch die Mitarbeitenden, wie in einigen Ländern bereits üblich, öffentlich publiziert wird. Für die Praxis heisst dies vor allem, dass ein spezifisches Auswertungsverfahren festgelegt und ein Wechsel vermieden werden sollte. Dabei sind die hier vorgestellten Erkenntnisse unbedingt zu berücksichtigen.

Prof. Dr. D. Schwappach, MPH

Leiter Forschung und Entwicklung von Patientensicherheit Schweiz und Dozent am Institut für Sozial und Präventivmedizin (ISPM), Universität Bern

Link zum Abstract:

<https://doi.org/10.1093/intqhc/mzx086>

Den Volltext können wir aus Copyright-Gründen leider nicht mit versenden.