



Elisabeth Schori

Elisabeth Schori, Berufs-
schullehrerin, MNSc cand.,
Interkantonale Schule für
Pflegerberufe
Landhausstrasse 19
CH-6340 Baar/Schweiz
e.schori@pflageschule-baar.ch

Dr. **Bernd Kersten**, Institut
für Psychologie, Universität
Bern/CH

Christoph Abderhalden,
MNSc, Forschungsstelle Pflege
und Pädagogik, Universitäre
Psychiatrische Dienste Bern/CH

Schlüsselwörter

Pflagesystem

Primary Nursing

Bezugspflege

Messinstrumente

Evaluation

Pflegeorganisationsform

Interraterreliabilität

Wie reliabel ist der Fragebogen zur Analyse der Pflegedokumentation (IzEP-Dok[©]) im Instrument zur Er- fassung von Pflegesystemen IzEP[©]?

Elisabeth Schori, Bernd Kersten, Christoph Abderhalden

Das „Instrument zur Erfassung von Pflegesystemen IzEP[©]“ dient der differenzierten Beschreibung des praktizierten Pflegesystems einer Station. Zu IzEP[©] gehört ein 30-Item-„Fragebogen zur Analyse der Pflegedokumentation IzEP-Dok“. Zur Untersuchung der Interrater-Reliabilität von IzEP-Dok beurteilten 5 RaterInnen 43 Pflegedokumentationen aus der Schweiz, aus Deutschland und aus Österreich. Die Übereinstimmung der 10 Raterpaare wurde für jedes Item mittels AC1- und Kappa-Statistik berechnet, für die Summenscores wurden Intraklassenkorrelationen (ICC) verwendet. IzEP-Dok hat sich in Bezug auf die meisten Items als reliables Einschätzungsinstrument erwiesen, in Bezug auf die Punktsummen war die Beurteilerübereinstimmung mittelmäßig. Zur Verbesserung der Interrater-Reliabilität sollten am Instrument und in der Anleitung Anpassungen vorgenommen werden.

Einleitung

Die Organisationsformen der Pflege haben sich im Laufe der Geschichte der beruflichen Pflege immer wieder verändert, beeinflusst durch jeweils vorherrschende Auffassungen von Pflege und Professionalität, durch betriebs- und arbeitspsychologische Erkenntnisse und Konzepte, sowie durch Veränderungen in der Pflegeausbildung (zum Beispiel Einführung verschiedener Ausbildungsniveaus) und durch Arbeitsmarktfaktoren (zum Beispiel Mangel an ausgebildeten Pflegenden).

In den letzten Jahrzehnten haben patientenorientierte Organisationsformen der Pflege wie Primary Nursing (PN) an Bedeutung gewonnen. Das in den 60er-Jahren in den USA entwickelte Primary Nursing stößt in den letzten Jahren auch im deutschsprachigen Raum auf immer größeres Interesse. Es finden Tagungen zum Thema statt, in Deutschland und in der Schweiz gibt es Netzwerke zu PN. Eine immer größere Zahl von Betrieben wählt PN als Pflegeorganisationsform.

Primary Nursing wird von Marram et al. (1979) definiert als „die Durchführung von umfassender, kontinuierlicher, koordinierter und individualisierter Pflege durch die Primary Nurse, die über die Autonomie, Rechenschaftspflicht und die Autorität verfügt, als verantwortliche Pflegeperson für ihre Patienten zu handeln“. Gemäß Manthey (2005) gibt es vier Schlüsselkonzepte, welche für PN konstituierend seien: Individuelle Verantwortung, Kontinuität der Pflege, direkte Kommunikation sowie die Kombination von Pflegeplanung und Pflegedurchführung in einer Hand.

In der Praxis unterscheiden sich das Verständnis von PN und der Grad der Umsetzung in verschiedenen Pflegeinstitutionen allerdings sehr, und auch die bisherigen Erkenntnisse über die Auswirkungen von PN sind nicht schlüssig. Besonders problematisch ist, dass das Konzept des untersuchten PN (d.h. der unabhängigen Variable, der Intervention) in den meisten Studien über Auswirkungen von PN zu wenig eindeutig definiert ist (Boeckler, 2006; Giovanetti, 1987; Thomas & Bond, 1991).

Entsprechend empfehlen Boumans und Landeweerd (1999) nicht nur Outcomes zu evaluieren, sondern den Grad der Umsetzung von PN besser operational zu erfassen. Auch Ryan und Logue (1998) empfehlen die Entwicklung von Methoden, um die Schlüsselemente von PN zu untersuchen. Die Studien sollten multiple Methoden von Datensammlungen und die Anwendung in diversen Settings – zum Beispiel Akutpsychiatrie, Langzeitpflege usw. umfassen. Zudem müssten die verwendeten Instrumente auf Validität und Reliabilität geprüft werden.

Abstract

The „Instrument zur Erfassung von Pflegesystemen IzEP©“ is a newly developed instrument to describe nursing care delivery systems, consisting of 9 modules. One of these modules is „IzEP-Dok©“, a 30-item-questionnaire to assess the documentation of the nursing process. To examine the interrater reliability of the questionnaire IzEP-Dok, 5 raters assessed 43 nursing records from Switzerland, Germany and Austria. Cohen’s Kappa and Gwets’ AC1-statistic were calculated to evaluate agreement on the level of each item, and intraclass-corellations (ICC) were used to determine agreement on the two sum scores. IzEP-Dok has proved to be reliable with respect to most items; the reliability of the two sum scores was moderate. In order to improve the interrater reliability, adaptations of the instrument and the instructions are recommended.

Es fehlt nach wie vor an einem wissenschaftlich abgestützten, im deutschsprachigen Raum breit einsetzbaren Instrument, um den Umsetzungsgrad des multidimensionalen Konstrukts Primary Nursing oder verwandter Pflegeorganisationsformen valide und reliabel zu prüfen und den Verantwortlichen zuverlässige und forschungsbasierte Entscheidungsgrundlagen zu liefern.

Das Instrument zur Erfassung von Pflegesystemen IzEP

Ein zur Schließung dieser Lücke literaturbasiert entwickeltes neues Instrument ist das „Instrument zur Erfassung von Pflegesystemen IzEP“, welches zurzeit psychometrisch getestet wird (Abderhalden et al., 2006). Mit dem Instrument werden fünf Merkmale der Pflegesysteme erfasst (Abbildung 1):

Merkmal	Definition
1: Orientierung an einem Konzept	Ausmaß, in dem sich die Pflege ideell und praktisch an einem Konzept orientiert
2: Verantwortung und Kontinuität	Ausmaß, in dem im Pflegesystem eine kontinuierliche Verantwortung für die Pflege individueller Patienten ideell und praktisch wahrnehmbar bei einzelnen Pflegenden liegt
3: Pflegeprozess	Ausmaß, in dem sich das Pflegesystem am Pflegeprozess orientiert
4: Kommunikation	Ausmaß, in dem die Zuständigkeiten für die Kommunikation mit und über Patienten klar geregelt sind und sich konzeptuell und praktisch am Ziel hoher personeller Kontinuität orientieren
5: Rollenverständnis	Ausmaß, in dem im Pflegesystem die Rolle der Pflegenden durch individuelle Verantwortung und Entscheidungsbefugnis für den gesamten Pflegeprozess gekennzeichnet ist

Abb. 1: Von IzEP erfasste Merkmale

Diese Merkmale nehmen Bezug zu den Schlüsselkonzepten von PN. Zusätzlich werden Merkmale der Station und des Personals erhoben (Kovariablen), die möglicherweise einen Einfluss auf die Wahl und die Umsetzung des Pflegeorganisationssystems haben.

IzEP besteht aus 9 Fragebögen, mit denen je Informationen aus einer Quelle erhoben werden (Befragung der Stationsleitung, der Vorgesetzten der Stationsleitung, von Pflegenden, Ärzten/Therapeuten, Patient, externen Stellen, Angehörige, Analysen der Pflegedokumentation und des Dienstplans). Der Fragebogen für die Pflegedokumentation (IzEP-Dok) besteht aus 30 Items, welche von Raterinnen eingeschätzt werden müssen. Die genaue Handhabung von IzEP-Dok ist in einer Anleitung festgehalten.

Die mit IzEP-Dok gewonnenen Informationen stellen Indikatoren für die Umsetzung des Pflegeprozesses dar (Merkmal 3 von IzEP). Dieses Merkmal hat zwei Dimensionen. Die Dimension A misst das Ausmaß, in dem die Schritte des Pflegeprozesses in der Praxis sichtbar und mit Einbezug der Patienten implementiert werden. Dimension B misst das Ausmaß, in dem in der Umsetzung der verschiedenen Schritte des Pflegeprozesses für individuelle Patienten eine hohe personelle Kontinuität erreicht wird. 25 Items beziehen sich auf Dimension A, also die Dokumentation der verschiedenen Schritte des Pflegeprozesses (Assessment, Problemformulierung, Pflegeplanung, Durchführung, Evaluation, Entlassungsplanung). Die mit diesen Items gewonnenen Informationen sollen zeigen, ob die einzelnen Schritte des Pflegeprozesses in der Dokumentation abgebildet sind. Die Antworten zu den entsprechenden Items liefern vorwiegend dichotome Nominaldaten (z. B. ja/nein, vorhanden/nicht vorhanden; vgl. Beispiel in Abbildung 2).

Keywords

- Nursing Care Delivery Systems
- Primary Nursing
- Research Instruments
- Measurement
- Evaluation
- Interraterreliability

ja	nein	Sind in der Pflegedokumentation pflegerelevante Informationen mit Hinweisen auf Risiken, Hilfebedarf und/oder Ressourcen bezogen auf physische und psychosoziale Aspekte erfasst?
<input type="checkbox"/>	<input type="checkbox"/>	Die Informationen sind vorhanden

Abb. 2: Beispiel für ein Item mit dichotomen Antwortkategorien

5 Items betreffen Dimension B, die personelle Kontinuität bei der Dokumentation der Pflegeprozess-Schritte. Die mit diesen Items gewonnenen Informationen sollen zeigen, ob die einzelnen Schritte des Pflegeprozesses von wenigen oder vielen verschiedenen Pflegenden dokumentiert worden sind. Beim Pflegesystem Bezugspflege/Primary Nursing müsste erwartet werden können, dass die Schritte des Pflegeprozesses mehrheitlich von der für den Patienten/die Patientin zuständigen pflegerischen Bezugsperson/der Primary Nurse dokumentiert werden. Die Antwortkategorien zu den entsprechenden Items liefern polytome Nominaldaten, das heißt, dass aus mehreren Antwortmöglichkeiten eine zutreffende gewählt werden muss. Abbildung 3 zeigt ein Beispiel.

Von wem wurde die Mehrzahl dieser pflegerelevanten Informationen dokumentiert?	
<input type="checkbox"/>	Die für den gesamten Aufenthalt des Patienten zuständige Pflegende oder ihre Vertretung
<input type="checkbox"/>	Eine/mehrere Pflegende <i>aus dem zuständigen Bereich/Gruppe</i>
<input type="checkbox"/>	Eine beliebige Pflegende
<input type="checkbox"/>	Eine Pflegende, die für diese Dokumentation bei allen Patienten zuständig ist
<input type="checkbox"/>	Nicht beurteilbar

Abb. 3: Beispiel für ein Item mit einer polytomen Antwortkategorie (eine zutreffende Antwort)

Den einzelnen Items ist ein Punktwert zugeordnet. Dieser fließt in die Gesamtpunktzahl, welche die zwei Dimensionen des Merkmals „Pflegeprozess“ von IzEP beschreibt. Die mit IzEP-Dok gewonnenen Daten ergeben 0 bis 26 Punktwerte für die Dimension A „Umsetzung der Pflegeprozess-Schritte“ und 0 bis 15 Punktwerte für die Dimension B „Personelle Kontinuität im Pflegeprozess“.

Bisher fehlen Untersuchungen zur Interraterreliabilität von IzEP-Dok. Die Feststellung der Zuverlässigkeit der Einschätzungen der Pflegedokumentationen ist eine Voraussetzung für die weitere Entwicklung und spätere Validierung des Instruments.

Fragestellung

- Wie reliabel ist die Beurteilung von Pflegedokumentationen (aus der Schweiz, Deutschland und Österreich) mit dem „Fragebogen zur Analyse der Pflegedokumentation IzEP-Dok“ im Instrument zur Erfassung von Pflegesystemen IzEP?
- Unterscheiden sich die Reliabilitätsparameter zwischen den drei Ländern?

Material und Methode Studienmaterial/Stichprobe

Als Studienmaterial wurden Kopien von anonymisierten Pflegedokumentationen aus schweizerischen, deutschen und österreichischen Krankenhäusern verwendet. Bei der Auswahl der Einrichtungen wurde eine möglichst große Vielfalt angestrebt bezüglich Fachbereichen und Art der Dokumentation. Diese Varianz zwischen den Untersuchungsobjekten sollte zur Generalisierbarkeit der Ergebnisse beitragen.

Die konkrete Auswahl der Dokumentationen und deren Anonymisierung erfolgten durch die Verantwortlichen in den beteiligten Einrichtungen. Es musste sich um Dokumentationen von Patienten handeln, die seit mindestens drei Tagen auf der Station gepflegt wurden. Die Institutionen wurden mit einer Studieninformation über die Absicht der Studie, die benötigten Unterlagen und das Vorgehen bei der Anonymisierung der Unterlagen unterrichtet. Ein Erhebungsblatt diente dazu, Angaben der Station über die Pflegeorganisationsform und Regelungen im Umgang mit der Patientendokumentation zu erfassen.

Raterinnen und Schulung

Als Raterinnen wurden 5 Pflegefachfrauen aus der Schweiz eingesetzt, die eine dreijährige Pflegeausbildung abgeschlossen und mindestens zwei Jahre Berufserfahrung hatten. 4 Raterinnen arbeiteten im Langzeitbereich und eine in einem Paraplegiker Zentrum. Alle wurden vom Arbeitgeber für die Mitarbeit an der Studie freigestellt.

Durchführung

Die Raterinnen wurden in einer Schulung dem Fragebogen von IzEP-Dok und den unterschiedlichen Dokumentationen vertraut gemacht. Es wurden Vorkehrungen zur Vermeidung der häufigsten Beurteilerfehler getroffen (halo-effekt, baseline error, Primacy/Recency Effekt, zentrale Tendenz). Alle 5 Raterinnen schätzten alle Dokumentationen unabhängig voneinander ein. Die Beurteilerinnen wurden angehalten, sich nicht untereinander auszutauschen. Die Forscherin machte punktuelle Kontrollen und stand bei Unklarheiten zur Verfügung.

In einem Prätest wurde das Verfahren an 7 Dokumentationen getestet und eine Auswertung vorgenommen. Aufgrund der Erfahrungen im Prätest und der Auswertung der Daten wurden der Fragebogen und die Anleitung für die Raterinnen optimiert.

Datenaufbereitung und -analyse

Die ausgefüllten Fragebogen wurden im Programm Epidata 3.02 zweifach eingegeben und systematisch verglichen. Zur Beurteilung der Raterübereinstimmung wurden für jedes Item pro Raterpaar a) Cohens Kappa und b) die AC1-Statistik, sowie für die zwei Punktwertsummen c) eine Intraklassenkorrelation (ICC) für einzelne Beurteiler (single measure) berechnet.

- a) *Cohens Kappa* ist das am häufigsten verwendete Maß zur Bestimmung der Güte der *zufalls-korrigierten* Übereinstimmung (Cohen, 1960; Wirtz & Caspar, 2002). Eine in der Literatur diskutierte Problematik von Kappa besteht darin, dass die Ausprägung von Kappa nicht nur von der Übereinstimmung, sondern auch von anderen Merkmalen des Raterverhaltens und der Datenverteilung beeinflusst wird – etwa der Häufigkeit, mit der die Rater einzelne Kategorien wählen (Feinstein & Cicchetti, 1990; Mayer, Nonn, Evers, & Osterbrink, 2004; Wirtz & Caspar, 2002). Die Kappa Statistik ist dann schwierig zu interpretieren, wenn die Grundrate der Antworten der einzelnen Kategorien nicht gleichverteilt ist (Ungleichverteilung der Randsummen). Insbesondere für den Fall einer vollständigen Beurteilerübereinstimmung ist Kappa nicht definiert. U.E. beträgt die Reliabilität dann 1 (oder nahe bei 1), und entsprechende Fälle wurden als perfekte Übereinstimmung gewertet. Cohens Kappa wird (bei gleicher prozentualer Beurteilerübereinstimmung) drastisch gesenkt, je häufiger das zu untersuchende Phänomen von beiden Beobachtern wahrgenommen wird. Nach dem so genannten 1. Paradox von Cohens Kappa wird in solchen Tabellen die Güte der Reliabilität stark unterschätzt (Mayer et al., 2004).
- b) Die *AC1-Statistik* wurde von Gwet (2001; 2002) als Alternative zu Kappa eingeführt. Sie soll zu einer valideren Einschätzung der Beurteilerübereinstimmung führen und vermeidet bei ungleich hohen Beobachterprävalenzen die Probleme mit Cohens Kappa. Für den Fall perfekter Übereinstimmung zwischen den Beurteilern beträgt AC1 gleich 1. Für den Fall gleich verteilter Beobachterprävalenzen fallen Kappa und AC1 identisch aus.
- c) Die Intraklassenkorrelation (ICC) wird zur Prüfung der Beurteilerübereinstimmung bei intervallskalierten Daten verwendet (Fleiss, Levin & Paik, 2003). Intraklassenkorrelationen können in verschiedenen Formen berechnet werden. Im vorliegenden Fall wurde die Übereinstimmung gleicher Werte (*absolute agreement*) im Hinblick auf einen einzelnen Rater (einen mittleren Rater; *single measure*) berechnet, da IzEP-Dok-Einschätzungen im praktischen Einsatz von jeweils einer Person vorgenommen werden.

Die Übereinstimmung der 10 Raterpaare wurde für jedes Item im Median der Kappa- und AC1-Werte zusammengefasst; als Streuungsmaß wurden die Quartile und der Interquartilabstand zwischen dem 1. und 3. Quartil berechnet.

Die Kappa-Werte wurden gemäß der Empfehlung von Landis und Koch (1977) interpretiert (vgl. Abbildung 3); dasselbe Schema wurde nach Rücksprache mit K. Gwet auch auf die AC1-Statistik angewendet (Gwet, 2006). Die ICC-Werte wurden nach demselben Schema interpretiert (Fleiss, Levin & Paik, 2003).

<i>Kappa-Wert/AC1-Statistik</i>	<i>Strength of Agreement</i>	<i>Übereinstimmung</i>
< 0.00	Poor	Keine
0.00 - 0.20	Slight	Sehr geringe
0.21 - 0.40	Fair	Geringe
0.41 - 0.60	Moderate	Mittlere
0.61 - 0.80	Substantial	Hohe
0.81 - 1.00	Almost perfect	Sehr hohe

Abb. 4: Interpretation der Maße für die Interrater-Übereinstimmung

Die Kappa- und ICC-Werte wurden mit dem Programm SPSS für Windows (Version 12.0) berechnet, die AC1-Statistik anhand der entsprechenden Formeln (Gwet, 2001, S. 49 ff) im Programm MS Excel.

Ethische Überlegungen

Um den Schutz persönlicher Daten sicherzustellen, wurden die Patientendokumentationen einer „faktischen Anonymisierung“ unterzogen. Diese wird in der Erklärung der European Science Foundation (ESF) für die Verwendung personenbezogener Daten zu Forschungszwecken folgendermaßen umschrieben: „Die Anonymität ist als gegeben anzusehen, wenn der Betroffene nur mit einem unverhältnismäßigen Aufwand von Zeit, Kosten und Personal identifiziert werden kann“ (ESF, 2002, S. 19). Alle Namen von Patienten, Angehörigen, Ärzten, Therapeuten, Jahreszahlen, Institutionsnamen, Ortsnamen, Berufe, Adressen und Telefonnummern wurden entfernt oder verändert, die Namen von Pflegenden wurden durch deren Kürzel ersetzt. Auf das Einholen einer informierten Zustimmung der betroffenen Personen wurde verzichtet. Dies ist gemäß ESF möglich, „ (...) wenn der Benutzer der Daten nicht in der Lage ist, den Betroffenen zu identifizieren“ (ESF, 2002).

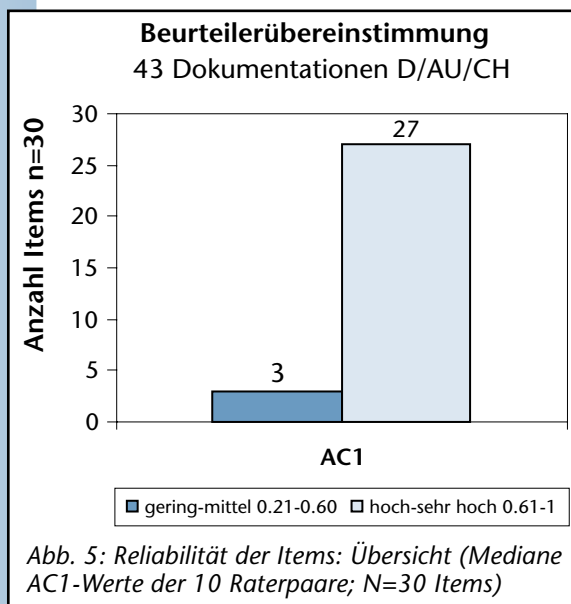
Ergebnisse

Beschreibung der Stichprobe

Untersucht wurden 43 Pflegedokumentationen aus 8 Krankenhäusern und 10 Stationen verschiedener Fachbereiche (vgl. Tabelle 1). Alle Dokumentationen wurden von allen 5 Raterinnen beurteilt. Somit lagen für die Auswertung 215 Einschätzungen von 5 Raterinnen vor.

Land	Anzahl Institutionen	Anzahl Stationen nach Fachgebieten	Anzahl Dokumentationen		Anzahl Einschätzungen	
Schweiz	4	2 Langzeitpflege	7	15	35	75
		1 Psychiatrie	4		20	
		1 Neuro/Rehabilitation	4		20	
Deutschland	3	1 Intensivstation	4	12	20	60
		1 Akut	4		20	
		1 Psychiatrie	4		20	
Österreich	1	2 Neurologie	11	16	55	80
		1 Intensivstation	5		25	
Total	8	10 Stationen	43			215

Tab. 1: Herkunft der verwendeten Dokumentationen



Die Dokumentationen waren sehr unterschiedlich gegliedert und gestaltet und der Detaillierungsgrad der Dokumentation variierte beträchtlich. Diese Unterschiede betreffen auch die pflegeprozessrelevanten Teile für die Pflegeanamnese, die Pflegebedarfserfassung, die Pflegeplanung und die Risikoeinschätzungen.

Interraterreliabilität

Mittels AC1-Statistik konnte die Übereinstimmung bei allen 30 Items anhand von je 10 Raterpaaren geprüft werden. Bei 27 Items (90%) liegen die Mediane der 10 AC1-Werte im Bereich von hoher bis sehr hoher, bei 3 Items im Bereich geringer bis mittlerer Übereinstimmung (Abbildung 5).

Die Kappa-Werte konnten aufgrund der Verteilung der Daten in 105 von 300 Fällen nicht

berechnet werden, bei 46 dieser Fälle lag eine perfekte Übereinstimmung vor, bei 59 eine sehr hohe prozentuale Übereinstimmung bzw. sehr hohe AC1-Werte. In 28 Fällen war ein negativer Kappa-Wert schwierig zu interpretieren, da die Übereinstimmung offensichtlich sehr hoch war.

Die Ergebnisdarstellung bezieht sich deshalb im Folgenden auf die AC1-Statistiken.

Die Übereinstimmung war bei den 25 Items aus den 11 Fragen zur **Umsetzung des Pflegeprozesses** etwas höher als bei den 5 Fragen, mit denen die **personelle Kontinuität** der Einträge in die Dokumentation eingeschätzt wird (Zusammenfassung in Tabelle 2, Details in Tabelle 3).

Interraterreliabilität		Inhaltlicher Aspekt		
		Umsetzung der Pflegeprozessschritte	Personelle Kontinuität im Pflegeprozess	Total IzEP-Dok
AC1-Wert		*	*	*
< 0.00	Keine Übereinstimmung	0	0	0 (0%)
0.00 - 0.20	Sehr geringe Übereinstimmung	0	0	0 (0%)
0.21 - 0.40	Geringe Übereinstimmung	0	0	0 (0%)
0.41 - 0.60	Mittlere Übereinstimmung	2	1	3 (10%)
0.61 - 0.80	Hohe Übereinstimmung	4	4	8 (26%)
0.81 - 1.00	Sehr hohe Übereinstimmung	19	0	19 (64%)
	Total	25	5	30 (100%)

Tab. 2: Interraterreliabilität der Items nach Dimensionen (Zusammenfassung; Zahlen sind Anzahl Items) * Anzahl Items

Übereinstimmung der Punktesummen

Die Intraklassenkorrelation (einzelne Maße) der Punktesummen betrug .466 (95%-Vertrauensintervall .307 - .625). bei der Dimension A „Ausmaß, in dem die Schritte des Pflegeprozesses in der Praxis sichtbar und mit Einbezug der Patienten implementiert werden“ und .440 (95%-VI .302 - .592) für die Dimension B „Ausmaß, in dem in der Umsetzung der verschiedenen Schritte des Pflegeprozesses für individuelle Patienten eine hohe personelle Kontinuität erreicht wird“. Diese Werte entsprechen einer mittleren Übereinstimmung.

Items mit der geringsten Übereinstimmung

Eine mittlere Übereinstimmung mit AC-1-Werten zwischen 0.54 und 0.58 erzielten die Raterinnen bei den Fragen, ob ein Zusammenhang der dokumentierten Wirkungen der Pflege mit den dokumentierten Problemen, Zielen und Maßnahmen erkennbar war (Item 23; AC1 0.58), ob pflegerische Aspekte der Entlassungsplanung dokumentiert waren (Item 24; AC1 0.56), und von wem die Mehrheit der Eintragungen zur Entlassungsplanung dokumentiert waren (Item 30; AC1 0.54).

Streuung der Übereinstimmung zwischen den 10 Raterpaaren

Bei 27 Items lagen die AC1-Werte der 10 Raterpaare nahe beieinander und der Interquartilabstand (1./3. Quartil) war kleiner als 0.2. Bei drei Items war die Differenz zwischen dem 2. und 4. Quartil größer als 0.2 (vgl. Tabelle 3, Spalte „Interquartil“). Folglich hatte wenigstens ein Teil der Raterinnen größere Schwierigkeiten die Frage zutreffend zu beantworten. Es handelt sich um die Fragen, ob Informationen zur Patientensicherheit dokumentiert sind (Item 13; 1./3.Quartil = 0.71/0.94, Interquartil = 0.23), um die Frage nach dem logischen Zusammenhang zwischen Problemen, Zielen und geplanten Maßnahmen (Item 20; 1./3. Quartil = 0.51/0.89, Interquartil = 0.38), und um die Frage, ob dokumentierte Wirkungen der Pflege in erkennbarem Zusammenhang mit dem Pflegeplan stehen (Item 23; 1./3.Quartil = 0.38 - 0.80; Interquartil = 0.42).

Vergleich der Übereinstimmung bei Dokumentationen von 7 der 30 Items

Abbildung 6 zeigt die Verteilung der Median-Werte des AC1 Schätzers über alle 30 Items. Der Anteil der hohen bis sehr hohen Übereinstimmung war bei den österreichischen Dokumentationen größer als bei den Dokumentationen aus der Schweiz und aus Deutschland (29 vs. 27 vs. 23 Items).

Item Nr.	Item	Alle (N=43)		Schweiz (N=15)	Deutschland (N=12)	Österreich (N=19)
		Median (1. - 3. Quartil)	Interquartil	Median	Median	Median
Umsetzung der Pflegeprozessschritte						
1	Assessmentinformationen vorhanden?	1.00 (1.00 - 1.00)	0.00	1.00	1.00	1.00
2	Info strukturiert dokumentiert?	0.98 (0.98 - 1.00)	0.02	0.93	1.00	1.00
3	Info an einem Ort dokumentiert?	0.89 (0.84 - 0.97)	0.13	0.87	1.00	1.00
4	Globale Einschätzung dokumentiert?	0.99 (0.98 - 1.00)	0.02	1.00	1.00	1.00
5	Info zu aktuellem Befinden dokumentiert?	0.97 (0.95 - 1.00)	0.05	1.00	0.91	1.00
6	Info zu Körperzustand dokumentiert?	0.98 (0.95 - 0.98)	0.03	0.93	1.00	1.00
7	Info zu Risikoabklärung dokumentiert?	0.96 (0.95 - 0.98)	0.03	0.93	0.91	1.00
8	Info zu Ernährung dokumentiert?	0.92 (0.79 - 0.93)	0.14	0.83	0.8	1.00
9	Info zu Ausscheidung dokumentiert?	0.91 (0.87 - 0.92)	0.05	0.69	1.00	1.00
10	Info zu Mobilität dokumentiert?	0.95 (0.91 - 0.95)	0.04	0.85	0.91	1.00
11	Info zu Körperpflege dokumentiert?	0.94 (0.90 - 0.97)	0.07	0.83	1.00	1.00
12	Info zu Kommunikation dokumentiert?	0.98 (0.98 - 0.99)	0.01	1.00	1.00	1.00
13	Info zu Sicherheit dokumentiert?	0.79 (0.71 - 0.94)	0.23	1.00	0.38	0.86
14	Wann erste Informationen dokumentiert?	0.88 (0.76 - 0.88)	0.12	0.79	0.83	0.87
15	Problem(e) dokumentiert?	1.00 (0.96 - 1.00)	0.04	1.00	1.00	1.00
16	Problemdefinition explizit?	0.82 (0.80 - 0.87)	0.07	0.90	0.47	1.00
17	Kein Plan dokumentiert?	1.00 (0.96 - 1.00)	0.04	1.00	1.00	1.00
18	Nur Ziele dokumentiert?	0.88 (0.78 - 0.98)	0.20	0.91	1.00	1.00
19	Maßnahmen dokumentiert?	0.98 (0.98 - 1.00)	0.02	0.92	1.00	1.00
20	Pflegeplan logisch zusammenhängend?	0.74 (0.51 - 0.89)	0.38	0.84	0.52	1.00
21	Durchführung Maßnahmen dokumentiert	0.68 (0.65 - 0.85)	0.20	0.50	0.50	1.00
22	Wirkungen der Pflege dokumentiert?	0.98 (0.95 - 0.98)	0.03	0.93	1.00	1.00
23	Zusammenhang dokumentierter Wirkungen mit Pflegeplan?	0.58 (0.38 - 0.80)	0.42	0.57	0.38	1.00
24	Entlassungsplanung dokumentiert?	0.56 (0.45 - 0.62)	0.17	0.63	0.10	0.58
25	Zusammenarbeit mit Patient dokumentiert?	0.72 (0.68 - 0.88)	0.2	0.85	0.66	0.82
Personelle Kontinuität im Pflegeprozess						
26	Von wem wurden Assessmentinformationen dokumentiert?	0.69 (0.65 - 0.77)	0.12	0.78	0.77	0.8
27	Von wem wurden Probleme dokumentiert?	0.68 (0.65 - 0.75)	0.10	0.71	0.77	0.7
28	Von wem wurde die Planung dokumentiert?	0.64 (0.59 - 0.74)	0.15	0.66	0.81	0.76
29	Von wem wurden Wirkungen dokumentiert?	0.62 (0.56 - 0.74)	0.18	0.67	0.65	0.8
30	Von wem wurde Entlassungsplanung dokumentiert?	0.54 (0.44 - 0.60)	0.16	0.47	0.38	0.62

Tab. 3: AC1-Werte aller Items und Ländervergleich

Literatur

Abderhalden, C.; Boeckler, U.; Dobrin Schippers; A., Feuchtinger, J.; Schaepe, C.; Schori, E.; et al. (2006): Ein Instrument zur Erfassung von Pflegesystemen (IzEP): Vorgehen bei der Instrumentenentwicklung. *PrInTerNet* (7/8), 420-424.

Boeckler, U. (2006): Primary Nursing einfach evaluieren? *PrInTerNet* (7/8), 404-419.

Boumans, N. P. G. & Landeweerd, J. A. (1999): Nurses' well-being in a primary nursing care setting in The Netherlands. *Scandinavian Journal of Caring Sciences*, 13(2), 116-122.

Cohen, J. (1960): A coefficient of Agreement for nominal scales. *Educational Psychological Measurement*, 10(1), 37-46.

ESF. (2002): Datenschutz in Wissenschaft und Forschung; 3.3.1 Erklärung der European Science Foundation zur Verwendung personenbezogener Daten zu Forschungszwecken, aus http://www.datenschutz-berlin.de/informat/dateien/mat_28.pdf am 13.05.2005

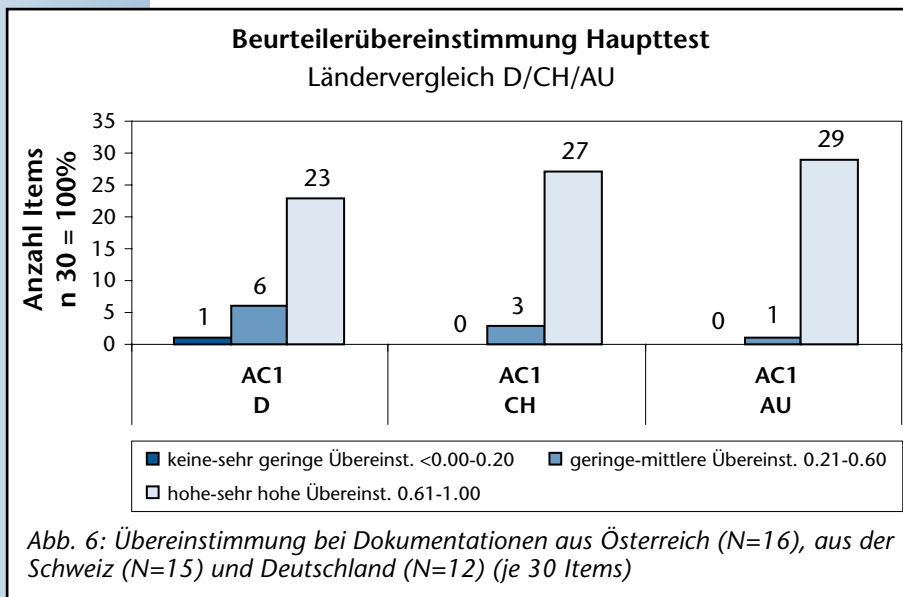
Feinstein, A. R. & Cicchetti, D. V. (1990): High Agreement but low Kappa: I. The Problems of two Paradoxes. *Journal of Clinical Epidemiology*, 43(6), 543-549.

Giovanetti, P. (1987): Gruppen- oder Einzelverantwortung in der Krankenpflege: Besteht ein Unterschied in der Qualität der Pflege? In Schweizer Berufsverband der Krankenschwestern und Krankenpfleger (Ed.), *Beurteilung der Pflegequalität: Referate der 4. Tagung über Forschung in der Krankenpflege des Schweizer Berufsverband der Krankenschwestern und Krankenpfleger (SBK)* (pp. 43-60). Bern: SBK, Schweizer Berufsverband der Krankenschwestern und Krankenpfleger.

Gwet, K. L. (2001): *Handbook of Inter-Rater Reliability, How to Estimate the Level of Agreement between Two or Multiple Raters*. Gaithersburg USA: STATAXIS Publishing Company.

Gwet, K. L. (2002): *Statistical Methods For Inter-Rater Reliability*

Die Übereinstimmung bei der Beurteilung der österreichischen Dokumentationen war bei 6 der 30 Items deutlich größer als bei den deutschen Dokumentationen (sehr hoch vs. mittel oder gering) und bei zwei dieser Items auch größer als bei den Schweizer Dokumentation (sehr hoch vs. mittel).



Diese 7 Items betreffen folgende Fragen:

- Problemdefinition explizit?
- Pflegeplan logisch zusammenhängend?
- Durchführung der Maßnahmen dokumentiert?
- Zusammenhang dokumentierter Wirkungen mit Pflegeplan?
- Info zu Sicherheit dokumentiert?
- Entlassungsplanung dokumentiert?
- Von wem wurde Entlassungsplanung dokumentiert?

Diskussion

Mittels der Einschätzung von 43 Pflegedokumentationen aus der Schweiz, Österreich und Deutschland durch 5 Raterinnen wurde die Interrater Reliabilität des Fragebogen zur Analyse der Pflegedokumentation IzEP-Dok getestet. Der Median der AC1 Werte lag bei 27 von 30 (90%) Items zwischen 0.61 und 1.00. Bei den zwei Punktesummen lagen die Intraklassenkorrelationen bei rund .45. Nach der Interpretation von Landis und Koch (1977) kann die Übereinstimmung bei den einzelnen Items als hoch bis sehr hoch, bei den Punktesummen als mittlere Beurteilerübereinstimmung umschrieben werden.

Bei 3 Items lagen die AC1 Werte unter 0.61, was einer mittleren Beurteilerübereinstimmung entspricht. Für die Beurteilung von Item 23 „erkennbarer Zusammenhang der dokumentierten Auswirkungen mit den genannten Problemen, Zielen und Maßnahmen“ (AC1 0.58) mussten die entsprechenden Informationen bei 9 Pflegedokumentationen infolge fehlender individueller Pflegeplanung an verschiedenen Stellen gesucht werden. Dieser große Aufwand führte offensichtlich dazu, dass ab und zu die Antwortvariante „nicht beurteilbar“ gewählt wurde. Schwierigkeiten mit diesem Item kommen auch in der größeren Streuung der Übereinstimmungsmaße zum Ausdruck (Distanz 1. - 3. Quartil > 0.2). Bei der Anpassung des Fragebogens muss das Streichen dieses Items diskutiert werden. Es ist auch fraglich, ob die Information dieses Items für die Erfassung des Pflegesystems relevant ist.

Mit dem Item zur Entlassungsplanung (24: AC1 0.56; 30: AC1 0,54) musste ein Element des Pflegeprozesses beurteilt werden, welches im Gegensatz zu den Elementen Assessment, Problemformulierung und Planung in der Praxis – und damit auch für die Raterinnen – relativ neu ist und in der Mehrzahl der Dokumentationen auch *noch nicht* als gesonderte Rubrik existiert. Die Beurteilung dieses Items wurde darüber hinaus dadurch erschwert, dass 9 Dokumentationen von Intensivstationen stammten. Hier würde die „Entlassungsplanung“ der Planung der Verlegung auf die Abteilung entsprechen. Dies muss im Fragebogen und im Raterleitfaden in diesem Sinn ergänzt werden.

Bei den Items 20 und 23 ist die Interquartildistanz größer als 0.2 (Range 1. - 3. Quartil 0.51-0.89 bzw. 0,38 - 0.8). Dies überrascht deshalb nicht, weil es sich um zwei Items handelt, bei denen der logische Zusammenhang von Problemen, Zielen und Maßnahmen (Item 20) bzw. der Zusammenhang dokumentierter Ergebnisse mit dem Plan beurteilt werden muss. Dabei besteht ein gewisser Ermessensspielraum, der vermutlich auch mit intensiverer Schulung nicht ganz reduziert werden kann.

Die Übereinstimmung bei der Beurteilung der österreichischen Dokumentationen war größer als bei den Schweizer Dokumentationen und deutlich größer als bei den deutschen Dokumentationen. Vermutlich werden die mit dem Instrument einzuschätzenden Elemente des Pflegeprozesses in der Schweiz und vor allem in Österreich erkennbarer und expliziter dokumentiert als in Deutschland. Dies ist für Österreich mit dem Gesundheits- und Krankenpflegegesetz (GuKG, 1.9.1997) zu erklären. Dort ist in § 5 festgelegt, dass die Angehörigen

- Assessment, Kappa Statistic is not Satisfactory for Assessing the Extent of Agreement Between Raters, aus <http://www.stataxis.com/files/articles/smirra1.pdf> am 30.10.2005
- Gwet, K. L. (2006): Persönliche Mitteilung; E-Mail vom 2. März 2006 an E. Schori.
- Kühne-Ponesch, S. & Smoliner, A. (2001): Die Entwicklung einer Fachsprache am Beispiel Österreichs: Die erste gesetzliche Regelung Europas. In N. Oud (Ed.), ACENDIO: Dritte europäische Konferenz der Organisation für gemeinsame europäische Pflegediagnosen, -interventionen und -ergebnisse (pp. 37-47). Bern: Huber.
- Landis, J. R. & Koch, G. G. (1977): The Measurement of Observer Agreement for Categorical Data. *Biometrics*, 33(1), 159-174.
- Manthey, M. (2005): *Primary Nursing, ein personenbezogenes Pflegesystem* (2 ed.). Bern: Huber.
- Marram, G.; Schlegel, M. & Bevis, O. (1979): *Primary Nursing* (2 ed.). St. Louis MO: C.V. Mosby.
- Mayer, H.; Nonn, C.; Evers, G. & Osterbrink, J. (2004): Qualitätskriterien von Assessmentinstrumenten Cohen's Kappa als Mass der Interrater-Reliabilität (Teil 1). *Pflege*, 17, 36-46.
- Ryan, A. & Logue, H. F. (1998): Developing an audit tool for primary nursing. *Journal of Clinical Nursing*(7), 417-423.
- Thomas, L. H. & Bond, S. (1991): Outcomes of nursing care: the case of primary nursing. *Int J Nurs Stud*, 28(4), 291-314.
- Wirtz, M. & Caspar, F. (2002): *Beurteilerübereinstimmung und Beurteilerreliabilität*. Göttingen: Hogrefe Verlag.

Kurzbiografie

Elisabeth Schori, dipl. Pflegefachfrau, Berufsschullehrerin, MNS(c), studiert Pflegewissenschaft im Studiengang Master in Nursing Science, Weiterbildungszentrum für Gesundheitsberufe Aarau CH/ Universität Maastricht NL. Sie ist seit 2004 Mitglied der Arbeitsgruppe zur Entwicklung eines Instrumentes zur Erfassung von Pflegesystemen (IzEP)

der Gesundheits- und Krankenpflegeberufe bei der Ausübung ihres Berufes die von ihnen gesetzten gesundheits- und krankenpflegerischen Maßnahmen zu dokumentieren haben und dass diese Dokumentation insbesondere die Pflegeanamnese, die Pflegediagnose, die Pflegeplanung und die Pflegemaßnahmen zu enthalten hat (Kühne-Ponesch & Smoliner, 2001). Entsprechend klare nationale Gesetzesvorgaben existieren in der Schweiz und in Deutschland nicht. In den Schweizer Dokumentationen sind die Pflegeprozesselemente möglicherweise deshalb besser erkennbar als in Deutschland, weil hier die aus dem angelsächsischen Raum stammenden Konzepte wie *nursing process*, *nursing diagnoses* etc. früher rezipiert wurden und die Praxis (und damit die Dokumentationsgewohnheiten) schon länger beeinflussen als in Deutschland. Zudem handelte es sich bei 5 der Schweizer Dokumentationen um Ausdrücke eines elektronischen Dokumentationssystems, was durch klare Strukturierung und gute Lesbarkeit die Beurteilung erleichterte.

Generell ist aufgefallen, dass die Beurteilung der Dokumentationen dann wesentlich leichter fiel, wenn Formulare zur individuellen Pflegeplanung vorhanden waren und/oder wenn mit NANDA Pflegediagnosen gearbeitet wurde.

Bei der Gewichtung der Ergebnisse dieser Studie müssen einige methodische Aspekte berücksichtigt werden. Das Ausmaß der Interraterreliabilität eines Instruments kann sowohl von den Raterinnen als auch von den bewerteten Objekten beeinflusst sein. Sowohl bei den Raterinnen als auch bei den beteiligten Einrichtungen und den untersuchten Dokumentationen handelt es sich um Gelegenheitsstichproben, bei denen nicht bekannt ist, inwiefern sie Merkmale aller potentiellen Raterinnen oder Dokumentationsformen repräsentieren. Eine Verallgemeinerung auf beliebige Raterinnen und beliebige Dokumentationen ist deshalb problematisch. Diese Einschränkung gilt allerdings ganz grundsätzlich für Studien zur Beurteilerübereinstimmung: Reliabilitätshinweise aus früheren Studien sollten nur als Orientierung und nicht als absolut gültige Schätzung verwendet werden (Wirtz & Caspar, 2002).

Insbesondere die Ergebnisse des Ländervergleichs dürfen nur vorsichtig interpretiert werden, da nicht bekannt ist, inwiefern die untersuchten Dokumentationen für die beteiligten Länder typisch oder repräsentativ sind. Die Dokumentationen aus Österreich stammen aus einem einzigen Betrieb.

Schlussfolgerung

Der Fragebogen zur Analyse der Pflegedokumentation IzEP-Dok hat sich in Bezug auf die meisten Items als reliables Einschätzungsinstrument erwiesen, in Bezug auf die Punktsummen war die Beurteilerübereinstimmung mittelmäßig. Die Ergebnisse des Ländervergleichs müssen infolge der geringen Anzahl mit Vorsicht interpretiert werden. Zur Verbesserung der Interrater-Reliabilität sollten am Instrument und in der Anleitung Anpassungen vorgenommen werden. Die Interraterreliabilität sollte im Rahmen erster Konstrukt-Validierungsstudien weiter getestet werden.