

Cluster-randomisierte Studien in der Pflegerwissenschaft

Oliver Kuß, Patrick Jahn, Petra Renz, Margarete Landenberger

23.1 Was sind cluster-randomisierte Studien?	304
23.2 Wann benötigt man cluster-randomisierte Studien?	305
23.3 Was ist anders in cluster-randomisierten Studien?	305
23.4 Ein Beispiel für eine cluster-randomisierte Studie	307
23.5 Empfehlungen für Cluster-randomisierte Studien	308
Literatur	309

Zusammenfassung

Es besteht weitgehende Einigkeit in der empirischen Gesundheits- und Pflegewissenschaft, dass Interventionen primär in randomisierten Studien geprüft werden sollen. Nur die Randomisierung garantiert eine gleichmäßige Verteilung aller bekannten und unbekannt Störgrößen und Risikofaktoren auf die Interventions- und Kontrollgruppe und erlaubt dadurch kausale Aussagen über Wirkungszusammenhänge.

Im Allgemeinen werden in randomisierten Studien die individuellen Probanden (oder Patienten) auf die Behandlungsgruppen zugeteilt. Es gibt jedoch Situationen, wo nicht individuelle Probanden, sondern ganze Gruppen (oder Cluster) von Probanden zufällig auf Interventions- und Kontrollgruppe aufgeteilt werden, wobei jedoch die Zielgrößen weiterhin an den individuellen Probanden erhoben werden. Man spricht in diesen Fällen von »cluster-randomisierten« Studien. Die Durchführung einer cluster-randomisierten Studie (anstatt einer individual-randomisierten) ist angezeigt, wenn 1) die Intervention auf natürliche Art und Weise in Clustern durchgeführt wird, 2) bei einer individual-randomisierten

Studie die Gefahr der Kontamination bestünde oder 3) eine cluster-randomisierte Studie kosteneffektiver als eine individual-randomisierte Studie ist.

Das Hauptproblem in cluster-randomisierten Studien stellt die Tatsache dar, dass die Ergebnisse von Probanden im selben Cluster nicht als unabhängig angenommen werden können. Dies hat Auswirkungen auf Design (insbesondere Fallzahlplanung), Auswertung und Berichterstattung von cluster-randomisierten Studien.

Wir beschreiben und diskutieren in dieser Arbeit einige der spezifischen Probleme von cluster-randomisierten Studien, sowohl allgemein als auch an einem konkreten Beispiel aus unserer Anwendung. Schließlich werden Empfehlungen gegeben, die empirisch arbeitenden Wissenschaftlern eine Hilfe bei der Planung, aber auch beim Lesen von cluster-randomisierten Studien sein sollen.

Abstract

It is commonly agreed in nursing research that interventions should be evaluated only in randomized trials. This is because only random allocation allows a valid measurement of the causal effect of the intervention under study.

In general, individual patients are assigned randomly to intervention and control group in randomized trials. However, there are situations where it is preferable to assign clusters of individuals to intervention groups. Cluster randomization is indicated when the intervention under study is 1) itself delivered to groups rather than individuals, 2) given to individuals but might affect others within that cluster, that is, contamination, or, 3) more cost-effective than individual randomization. Such trials are known today as »cluster-randomized« or »group-randomized« trials.

The main consequence of a cluster-randomized design is that, unlike individually randomized trials, the outcome for each participant cannot be assumed to be independent from that of any other participant in the same cluster. Instead, participants in the same cluster are more likely to have similar outcomes. This lack of independence influences design, analysis and reporting of cluster randomized trials.

We explain and illustrate the specific problems of such trials by using an example from our own research and give some recommendations for conducting and reading cluster-randomized trials.

Schlagworte

- Pflegewissenschaft
 - Cluster-randomisiert
 - Intra-Cluster-Korrelation
 - Gemischte Modelle
 - nursing research
 - cluster-randomized
 - group-randomized
 - intra-cluster-correlation
 - mixed models
-

23.1 Was sind cluster-randomisierte Studien?

Es besteht weitgehende Einigkeit in der Gesundheits- und Pflegewissenschaft sowie in allen anderen empirischen Wissenschaften, dass Interventionen oder Behandlungen primär in randomisierten Studien geprüft werden sollen. Nur die Randomisierung garantiert eine gleichmäßige Verteilung aller bekannten und unbekannt Störgrößen und Risikofaktoren auf die Interventions- und Kontrollgruppe. Unterschiede in den Zielgrößen, die im Studienverlauf in diesen Gruppen auftreten, sind daher ausschließlich auf die Intervention und nicht auf vorher bestehende Unterschiede zwischen Interventions- und Kontrollgruppe zurückzuführen. Die randomisierte Studie ist in diesem Sinne das einzige Studiendesign in der empirischen Forschung, das, ohne zusätzliche Annahmen und bis auf zufällig entstandene Imbalancen, eine kausale (im Sinne einer echten Ursache-Wirkungsbeziehung) Aussage über die Wirksamkeit von Interventionen erlaubt.

Im Allgemeinen werden in randomisierten Studien die individuellen Probanden (oder Patienten) auf die Behandlungsgruppen zugeteilt. Es gibt jedoch Situationen, in denen nicht individuelle Probanden, sondern ganze Gruppen (oder Cluster) von Probanden zufällig auf Interventions- und Kontrollgruppe aufgeteilt werden, wobei jedoch die Zielgrößen weiterhin an den individuellen Probanden erhoben werden. Man spricht in diesen Fällen von »Cluster-Randomisierung« oder »cluster-randomisierten Studien«.

Cluster-randomisierte Studien werden immer häufiger durchgeführt, auch in der Pflegewissenschaft¹⁻³. So werden z.B. in 3 von 5 Studien in der 2. Förderphase des Pflegeforschungsverbund Mitte-Süd cluster-randomisierte Studiendesigns eingesetzt. Es gibt daneben bereits eine Erweiterung der CONSORT-Richtlinie⁴ zur Darstellung von cluster-randomisierten Studien. Neben den grundlegenden Lehrbüchern^{5,6} wurde inzwischen eine Reihe von Übersichtsartikeln in wissenschaftlichen Journalen⁷⁻¹¹ veröffentlicht, die auch Nicht-Statistikern einen Zugang zu diesen Studien ermöglichen sollen.

23.2 Wann benötigt man cluster-randomisierte Studien?

Individual-randomisierte Studien sind cluster-randomisierten Studien im Allgemeinen vorzuziehen, da in ersteren eher eine Strukturgleichheit zwischen Interventions- und Kontrollgruppe auf Probandenebene erreicht werden kann.

Es gibt jedoch mindestens drei Situationen, wo eine cluster-randomisierte Studie einer individual-randomisierten Studie vorgezogen werden sollte^{11–13}.

Zum ersten wird das überall der Fall sein, wo die Intervention in natürlicher Art und Weise bereits auf Gruppen angewandt wird, bspw. wenn der Effekt einer neuen Organisationsform (z.B. gemeinsame vs. individuell eingenommene Mahlzeiten) in einem Pflegeheim geprüft werden soll. Hier werden sich notgedrungen alle Bewohner des Pflegeheims in die heimweit eingeführte Organisationsform eingliedern.

Zum zweiten wird man überall dort eine Cluster-Randomisierung durchführen wollen, wo es zu einer Kontamination des Interventionseffektes kommen könnte, weil sich Probanden aus der Interventionsgruppe mit den aus der Kontrollgruppe austauschen und so Teile der Intervention auch in die Kontrollgruppe tragen könnten. Hiervon sind meistens Interventionen gefährdet, die als Schulungsmaßnahme durchgeführt werden. Wird z.B. der Effekt eines Programms zur Prävention von Stürzen in Pflegeheimen überprüft, so sollten alle Einwohner eines Pflegeheimes derselben Gruppe (Kontrolle oder Intervention) angehören. Andernfalls bestünde die Gefahr, dass Bewohner, die in die Interventionsgruppe randomisiert worden sind, sich mit Bewohnern der Kontrollgruppe über Einzelheiten des Präventionsprogramms austauschen.

Zum dritten kann es auch wesentlich effektiver sein, eine Intervention an einer Gruppe zu prüfen, weil eine Intervention an jedem einzelnen Probanden zu aufwändig wäre. Auch hier kann wieder eine Schulungsintervention als Beispiel dienen. Es wäre sehr aufwändig, jeden einzelnen Probanden für die Studie individuell zu schulen. Tatsächlich wird es in der Regel kosteneffektiver sein, die Schulung in Gruppen durchzuführen.

23.3 Was ist anders in cluster-randomisierten Studien?

Das Hauptproblem in cluster-randomisierten Studien stellt die Tatsache dar, dass die Ergebnisse von Probanden im selben Cluster nicht als unabhängig angenommen werden können. Wenn die Probanden vom selben Lehrenden unterrichtet oder auf derselben Station versorgt worden sind, dann werden sich die Einflüsse von Lehrenden oder Stationen auch auf die Probanden auswirken. Zwei Probanden eines »guten« Lehrenden werden eher beide ein »gutes« Ergebnis haben, zwei Probanden eines »schlechten« Lehrenden eher ein »schlechtes«. Auf verschiedenen Stationen werden Patienten mit unterschiedlicher Krankheitsschwere versorgt. Zwei Schwerkranke auf einer onkologischen Station werden unter Umständen ähnlich schwierig auf eine Intervention reagieren und eher ein »schlechtes« Ergebnis haben, im Vergleich zu zwei Probanden auf einer Station, auf der weniger schwere Krankheiten behandelt werden und die dann eher ein »gutes« Ergebnis erreichen.

Die dadurch entstehende Korrelation zwischen den Probanden oder Patienten, die so genannte Intra-Cluster-Korrelation (ICC), muss sowohl bei der Planung als auch bei der Auswertung einer cluster-randomisierten Studie berücksichtigt werden. Da die herkömmlichen statistischen Verfahren in der Regel eine Unabhängigkeit der Daten verlangen, sind die in cluster-randomisierten Studien anzuwendenden Verfahren von höherer Komplexität. In der Regel werden cluster-randomisierte Studien mit Hilfe der gemischten Modelle ausgewertet^{14,15}. Hier wird das zugrunde liegende statistische Modell um einen zufälligen Faktor für die Cluster erweitert, der die durch das Design entstandene Intra-Cluster-Korrelation mit berücksichtigen kann. Diese Modelle sind inzwischen relativ weit entwickelt, sie können mit Standardsoftware (z.B. SAS, SPSS oder R) geschätzt werden und auch Erweiterungen auf nicht-stetige Zielgrößen oder noch komplexere Abhängigkeiten (z.B. zusätzlich zur Clusterung noch wiederholte Beobachtungen im Zeitverlauf) können modelliert werden. Von der immer wieder anzutreffenden Praxis, die Informationen der Probanden innerhalb eines Clusters zusammenzufassen und die weitere Auswertung nur mit dieser aggregierten Information durchzuführen ist abzuraten. Erstens findet so ein unnötiger Informationsverlust statt, des weiteren wird die Unsicherheit bei der Schätzung der aggregierten Zielgröße im weiteren Verlauf der Auswertung ignoriert, so dass Konfidenzintervalle und p-Werte zu klein werden und eine nicht gegebene Sicherheit vorgaukeln.

Nach wie vor eine große Herausforderung ist die Fallzahlplanung in cluster-randomisierten Studien. Hier benötigt man zu den herkömmlichen Angaben, die man auch für individual-randomisierte Studien benötigt, eine zusätzliche Annahme über die erwartete Intra-Cluster-Korrelation. Diese Information ist auch deshalb wichtig, weil sich der Wert der erwarteten ICC unmittelbar und unter Umständen massiv auf die benötigte Fallzahl auswirkt. Dabei gilt, dass je größer die ICC, desto größer auch die benötigte Fallzahl ist. Glücklicherweise sind die ICC, die in pflegewissenschaftlichen Studien zu erwarten sind, eher klein. Eine Analyse von 1.039 Intra-Cluster-Korrelationen im verwandten Bereich der Allgemeinmedizin hatte einen medianen ICC von 0,01 ergeben¹⁶. Darüber hinaus ist die benötigte Fallzahl auch von der Anzahl der Cluster abhängig: 10 Probanden in 100 Clustern führen zu einer größeren statistischen Power als 100 Probanden in 10 Clustern. Insofern ist bei der Planung einer cluster-randomisierten Studie auch darauf zu achten, dass möglichst viele verschiedene Cluster rekrutiert werden. Eine cluster-randomisierte Studie mit nur zwei Clustern ist sinnlos, da der Effekt des Clusters nicht vom Interventionseffekt unterschieden werden kann.

Die in cluster-randomisierten Studien häufig anzutreffenden kleinen Clusterzahlen führen zu einem weiteren Problem, nämlich zu Imbalancen wichtiger prognostischer oder Risikofaktoren zwischen Interventions- und Kontrollgruppen, wenn der Randomisierung nicht ausreichend Cluster unterworfen waren, um für eine Strukturgleichheit zu sorgen. Eine Lösung dieses Problems kann, bereits in der Design-Phase der Studie, ein Matchen von Clustern mit ähnlichen Eigenschaften sein, so dass später nur noch innerhalb der Matching-Strata randomisiert wird. Eine zweite Lösung wäre, für eventuell aufgetretene Imbalancen mit Hilfe statistischer Modelle zu adjustieren, d.h. die imbalanzierten Merkmale als Kovariablen ins Auswertungsmodell mit aufzunehmen. In diesem Fall sollte

allerdings dieser Selektionsalgorithmus für die Kovariablen bereits vorher spezifiziert worden sein.

23.4 Ein Beispiel für eine cluster-randomisierte Studie

Die Autoren haben in der 1. Förderphase des Pflegeforschungsverbund Mitte-Süd eine cluster-randomisierte Studie durchgeführt. Wir berichten hier über einige methodische Spezifika dieser Studie, für inhaltliche Ergebnisse verweisen wir auf die begleitenden Publikationen¹⁷.

Ziel dieser Studie war die Reduktion von Chemotherapie-induzierter Anorexie, Nausea und Emesis (ANE) durch eine strukturierte onkologisch-fachpflegerische Intervention. Diese Intervention war in einer cluster-randomisierten Studie in 14 onkologischen Stationen (diese repräsentieren die Cluster) an zwei deutschen Universitätskliniken (Halle (Saale), München) geprüft worden. Die Intervention bestand dabei aus 4 Modulen (1. Informationsbroschüre/Beratungsgespräch, 2. Pflegerische Mitwirkung bei der Antiemese-Medikation, 3. Ernährungsberatung, 4. Pflegerische/ interdisziplinäre Entspannung) und wurde von den Pflegekräften der Stationen in der Interventionsgruppe zusätzlich zur kliniküblichen Pflege durchgeführt. Auf den Stationen der Kontrollgruppe wurde nur die klinikübliche Pflege (inkl. der herkömmlichen Antiemese-Therapie) durchgeführt. Die Hauptzielgröße war dabei die Reduktion der ANE-Symptomatik beim individuellen Patienten, gemessen mit dem CTCAE-Score¹⁸. Nebenzielgrößen betrafen dabei das Wissen der Patienten zu Nebenwirkungen, zur Selbstpflege, die Selbstpflegekompetenz und die gesundheitsbezogene Lebensqualität.

Die Gründe für die Durchführung einer cluster-randomisierten (anstelle einer individual-randomisierten) Studie sind hier offensichtlich. Um eine Kontamination auf Patientenebene zu vermeiden (da sich Patienten ein und derselben Station über Inhalte des Interventionsprogramms austauschen könnten), müssen alle Patienten auf einer Station derselben Behandlungsgruppe angehören. Daneben war es aber auch nötig, alle Pflegenden auf einer Station in die gleiche Behandlungsgruppe einzuteilen, denn nur so war eine kontinuierliche Durchführung der Intervention im täglichen, schichtbedingten Wechsel der Pflegekräfte gewährleistet. Eine Individual-Randomisierung hätte zudem dazu geführt, dass Patienten in der Kontrollgruppe u. U. von Pflegenden betreut worden wären, die parallel auch Patienten aus der Interventionsgruppe betreuen. Insofern hätte auch die Gefahr einer Kontamination der Patienten der Kontrollgruppe durch die (in der Intervention geschulten) Pflegekräfte bestanden.

Die Fallzahlplanung für die Hauptzielgröße basierte auf einer angenommenen ICC von 0,05. Daraus resultierte eine benötigte Fallzahl von 200 Patienten (100 Patienten in der Kontroll- und 100 in der Interventionsgruppe). Zum Vergleich, der zugrunde gelegte Effekt hätte bei Individual-Randomisierung bereits mit 140 Patienten (70 Patienten in der Kontroll- und 70 in der Interventionsgruppe) entdeckt werden können. Man hätte bei Individual-Randomisierung also 60 Patienten »einsparen« können.

Wir stellen hier exemplarisch die Ergebnisse für die Nebenzielgröße »Wissen über Nebenwirkungen« (ohne Adjustierung für imbalanzierte Baseline-Merkmale) dar. Diese

war mit Hilfe einer 100mm Visual-Analog-Skala mit den Extremwerten 0=»unzureichend« bis 100=»sehr gut« am 8. Tag des zweiten Chemotherapiezyklus erhoben worden.

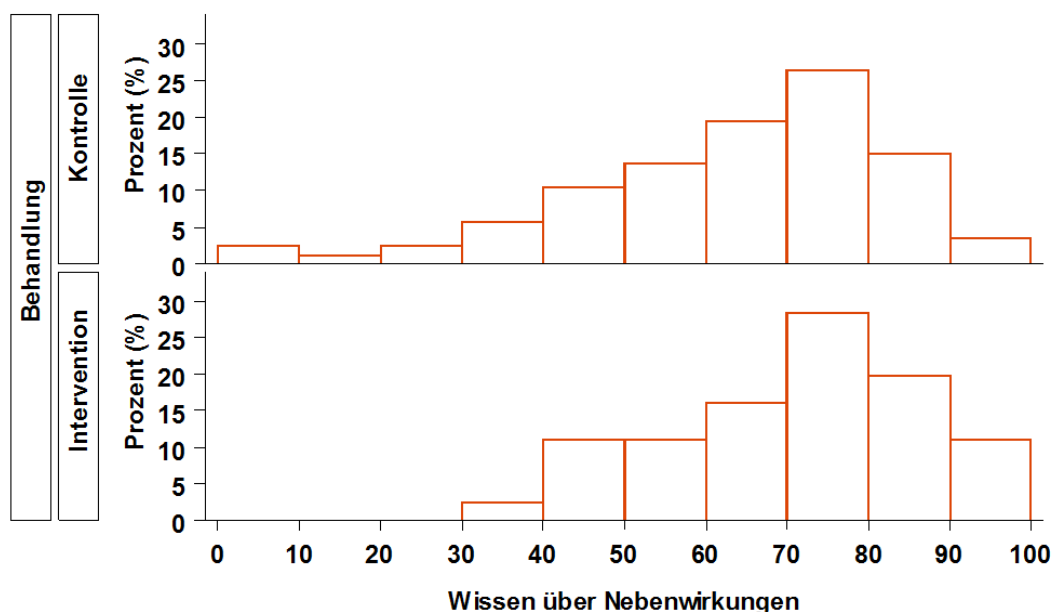


Abbildung 23.1: Deskriptive Darstellung (Histogramme) der Nebenzielgröße »Wissen über Nebenwirkungen« (100mm VAS, 0=»unzureichend« bis 100=»sehr gut«) in den beiden Behandlungsgruppen der P4-Studie¹⁷

Eine rein deskriptive Darstellung mit Hilfe zweier Histogramme (Abbildung 23.1) über alle Patienten ohne Berücksichtigung der Clusterung zeigt einen leichten Vorteil der Intervention. Patienten in der Interventionsgruppe geben dabei ein mittleres Wissen von 71,7 mm an, die in der Kontrollgruppe von 64,2 mm. Die Differenz von 7,5 mm mit einem 95%-Konfidenzintervall von [2,2; 12,8] wäre dabei mit einem herkömmlichen t-Test auch signifikant ($p=0,006$) von Null verschieden. Diese Auswertung mit einem herkömmlichen t-Test berücksichtigt jedoch *nicht* die Clusterung und eine angemessene Analyse mit einem gemischten Modell ergibt einen Effekt von 7,1 mm [95%-KI: -2,0; 16,2], der jetzt nicht mehr signifikant ($p=0,11$) ist. Die geschätzte ICC ergibt sich dabei zu 0,07 [95%-KI: -0,05; 0,19]. Der Interventionseffekt wird hier also ohne Berücksichtigung der Clusterung überschätzt.

23.5 Empfehlungen für Cluster-randomisierte Studien

Die folgenden Empfehlungen (in Anlehnung an Wears¹¹ und Ukoumunne et al.¹²) sollen empirisch arbeitenden Wissenschaftlern eine Hilfe bei der Planung, aber auch beim Lesen von cluster-randomisierten Studien sein.

1. Individual-randomisierte Studien sind cluster-randomisierten Studien im Allgemeinen vorzuziehen. Überlegen Sie also, ob eine cluster-randomisierte Studie gerecht-

fertigt ist und nicht besser individual randomisiert werden sollte. Falls doch eine cluster-randomisierte Studie durchgeführt werden muss, machen Sie die Gründe dafür explizit. Gute Gründe könnten z.B. sein:

- Die Intervention wird auf natürliche Art und Weise in Clustern durchgeführt.
 - Bei einer individual-randomisierten Studie bestünde die Gefahr der Kontamination.
 - Eine cluster-randomisierte Studie ist kosteneffektiver als eine individual-randomisierte Studie.
2. Berücksichtigen Sie im gesamten Verlauf der Studie das Prinzip der Cluster-Randomisierung. Dies betrifft insbesondere für die Fallzahlplanung, die Auswertungsmethodik und die Berichterstattung zu.
 3. Schließen sie genügend Cluster ein. Die Power einer cluster-randomisierten Studie hängt stärker von der Anzahl der Cluster als von der Gesamtprobandenzahl ab. Als Daumenregel: Weniger als insgesamt 8-10 Cluster sollten nicht eingeschlossen werden.
 4. Berichten Sie die Werte der geschätzten ICC. Sie erleichtern damit die Planung von zukünftigen Studien.
 5. Überlegen Sie, ob Cluster vor der Randomisierung gematcht werden sollten. Dies ist vor allem dann wichtig, wenn insgesamt nur wenige oder sehr unterschiedliche Cluster zur Verfügung stehen.

Literatur

1. Fossey J, Ballard C, Juszczak E et al. Effect of enhanced psychosocial care on antipsychotic use in nursing home residents with severe dementia: cluster randomised trial. *BMJ* 2006 April 1;332(7544):756-61.
2. Kovacs F, Abaira V, Santos S et al. A comparison of two short education programs for improving low back pain-related disability in the elderly: a cluster randomized controlled trial. *Spine* 2007 May 1;32(10):1053-9.
3. Cox H, Puffer S, Morton V et al. Educating nursing home staff on fracture prevention: a cluster randomised trial. *Age Ageing* 2008 March;37(2):167-72.
4. Campbell MK, Elbourne DR, Altman DG. CONSORT statement: extension to cluster randomised trials. *BMJ* 2004 March 20;328(7441):702-8.
5. Murray DM. *Design and analysis of group-randomized trials*. New York: Oxford University Press; 1998.
6. Donner A, Klar N. *Design and analysis of cluster randomization trials in health research*. London: Arnold; 2000.

7. Bland JM, Kerry SM. Statistics notes. Trials randomised in clusters. *BMJ* 1997 September 6;315(7108):600.
8. Kerry SM, Bland JM. Analysis of a trial randomised in clusters. *BMJ* 1998 January 3;316(7124):54.
9. Kerry SM, Bland JM. The intraclass correlation coefficient in cluster randomisation. *BMJ* 1998 May 9;316(7142):1455.
10. Kerry SM, Bland JM. Sample size in cluster randomisation. *BMJ* 1998 February 14;316(7130):549.
11. Wears RL. Advanced statistics: statistical methods for analyzing cluster and cluster-randomized data. *Acad Emerg Med* 2002 April;9(4):330-41.
12. Ukoumunne OC, Gulliford MC, Chinn S, Sterne JA, Burney PG, Donner A. Methods in health service research. Evaluation of health interventions at area and organisation level. *BMJ* 1999 August 7;319(7206):376-9.
13. Medical Research Council. Cluster randomised trials: Methodological and ethical considerations. http://open.mrc.gm/Statsdatamgt/stats&datam_files/MRCguide_cluster.pdf. 2002.
14. Brown H, Prescott R. *Applied mixed models in medicine*. 2nd ed. Chichester: Wiley; 2006.
15. Kuss O, Watzke S. [Proper handling of correlated data in rehabilitation research]. *Rehabilitation (Stuttg)* 2005 December;44(6):367-72.
16. Adams G, Gulliford MC, Ukoumunne OC, Eldridge S, Chinn S, Campbell MJ. Patterns of intra-cluster correlation from primary care research to inform study design and analysis. *J Clin Epidemiol* 2004 August;57(8):785-94.
17. Renz P, Jahn P, Landenberger M et al. Reduktion von chemotherapie-induzierter Anorexia, Nausea und Emesis durch eine strukturierte Pflegeintervention. In: Schaeffer D, Behrens J, Görres S, editors. *Optimierung und Evidenzbasierung pflegerischen Handelns*. Weinheim: Juventa; 2009. p. 189-210.
18. NCI-CTEP. CTCAE Common Terminology Criteria for adverse events (Version 3.0). <http://ctep.cancer.gov/reporting/ctc.html>. 2003.