

- [9] Borkowski A, Lee DH, Sydnor DL, Johnson RJ, Rabinovitch A, Moore GW. Intranet-based quality improvement documentation at the Veterans Affairs Maryland Health Care System. *Mod Pathol.* 2001 Jan; 14(1): 1–5.

Abstraktnummer 35/1

Randomisierte Interventionsstudien zur Evaluation neuer Medien in der medizinischen Lehre

Schulz S, Auhuber T, Klar R

Universitätsklinikum Freiburg, Medizinische Informatik, Freiburg

Die Methodik der randomisierten, kontrollierten Studie wird oft gleichgesetzt mit patientenorientierter klinischer Forschung, wobei zumeist vergessen wird, dass beispielsweise die Technik der Randomisierung zuerst bei landwirtschaftlichen Experimenten verwendet wurde, und dass Psychologie, Wirtschafts- und Sozialwissenschaften ebenfalls für sog. Ex Post-Analysen ein vergleichbares Methodeninventar entwickelt haben.

Zweifelsohne werden die meisten wissenschaftlichen randomisierten Interventionsstudien im Bereich der Medizin durchgeführt, mit bisher ca. 170 000 Studien laut MEDLINE¹). Im nichtmedizinischen Bereich dominiert die empirische Kriminologie [3] (6000) vor der empirischen Sozialforschung und der Psychologie (je 3000), während die empirische Bildungsforschung mit bisher weniger als tausend randomisierten Studien weit abgeschlagen bleibt [2].

Dies überrascht, wurde doch bereits in den sechziger Jahren gefordert, Reformen im Bildungsbereich durch experimentelle Begleitforschung auf ihre Zweckmäßigkeit zu überprüfen. Tatsächlich beschränkt man sich bisher meist darauf, Struktur- und Prozessqualität zu messen, während Aussagen zur Ergebnisqualität aufgrund fehlender Beobachtungs- und Strukturgleichheit der Gruppen weitgehend spekulativ sind. So resultiert eine Fülle von Forschungsergebnissen mit eher bescheidener Evidenz, und speziell für hypermediale Lernsysteme wird „im Land der Null-Hypothesen“ eine „Nicht-Evaluierbarkeit“ postuliert [7]. Selbst in den USA, wo weitaus am meisten Bildungsforschung betrieben wird, wurde im Jahre 2000 gerade mal ein einziges Forschungsvorhaben mit einem randomisierten Studiendesign öffentlich gefördert [4, 5].

Eine Medline-Recherche erbrachte 38 randomisierte Studien zu elektronischen Medien in der medizinischen Aus- und Weiterbildung. Methodisch steht unter den wenigen randomisierten Vergleichsstudien das sogenannte Crossover-Design im Vordergrund, welches an einer eigenen Studie kurz skizziert werden soll. Gegenstand unserer Untersuchung war hierbei ein interaktiver Pathologieatlas, der kursbegleitend zum Selbstlernen genutzt wird und das Prüfungswissen des Fachs allg. Pathologie umfasst. Er enthält Tausende mikroskopischer Bilder in unterschiedlicher Vergrößerung, die mit kurzen Lerntexten versehen sind.

Um zu prüfen, ob dieses Lernprogramm dem Selbstlernen mit Lehrbüchern und Atlanten überlegen ist, teilten wir 72 Probanden aus dem Kurs der allgemeinen Pathologie durch Randomisierung in zwei Gruppen ein. Gruppe I erarbeitete einen Themenkomplex A mit Hilfe des Programms und den Themenkomplex B mit Hilfe der Bücher. Gruppe II hingegen verwendete Bücher für das Thema A, und das Programm zur Erarbeitung von B.

¹) <http://www.ncbi.nlm.nih.gov/PubMed/>

Im Anschluss mussten alle Probanden je drei unbekannte Präparate pro Themenbereich am Mikroskop identifizieren. Als Ergebnis zeigte sich ein arithmetisches Mittel von 1,85 richtig beantworteten Fragen beim Lernen mit Buch, sowie von 1,97 beim Lernen mit Computer, so dass die Nullhypothese, nämlich die der Gleichwertigkeit beider Lernformen, nicht abgelehnt werden konnte [1].

Ein wesentlicher Vorteil eines Crossover-Versuchsplans [6] liegt darin, dass jeder Proband sowohl Test- als auch Kontrollperson ist. Dadurch kommt man mit kleineren Stichproben aus, und die Forderung nach Strukturgleichheit zwischen den Gruppen ist in nahezu idealer Weise verwirklicht. Weiterhin besteht Chancengleichheit zwischen beiden Gruppen, selbst unter der Annahme, dass mit den beiden Lehrmethoden deutlich unterschiedliche Ergebnisse erzielt werden. Somit können derartige Studien im Rahmen von Routineveranstaltungen realisiert werden. Berücksichtigt werden müssen – ebenso wie bei Crossover – Studien im klinischen Bereich, Perioden- und Residualeffekte. Die größte Schwierigkeit bei Bildungsstudien ist jedoch die Auswahl der Zielparameter, sowie die Beurteilung ihrer Validität. Sicherlich empfiehlt es sich, im Rahmen von Studien diejenigen Mechanismen zur Leistungsbewertung zu verwenden, die auch sonst Anwendung finden, also Klausuren, sowie praktische und mündliche Prüfungen. Ein weiteres Problem ist, dass eine Verblindung der Gruppen nicht möglich ist. Unabdingbar ist dagegen die Verblindung bei der Erfolgskontrolle einerseits im Hinblick auf die Gruppenzugehörigkeit der Prüflinge, und andererseits hinsichtlich der Kenntnis von Details der miteinander konkurrierenden Lernszenarien. Keinesfalls dürfen also die Prüfer in die Herstellung der verwendeten Lernmittel involviert sein. Da Lernerfolg auch darin besteht, die erlernten Inhalte möglichst lange im Gedächtnis zu behalten, ist weiterhin zu fordern, die Nachhaltigkeit des Wissens durch mindestens zwei Erfolgskontrollen in zeitlichem Abstand zu bewerten.

Trotz aller methodischer und organisatorischer Probleme sollte der Methodenstandard klinischer Forschung auch von der Bildungsforschung erfüllt werden. Dieser Standpunkt wird auch unterstützt von der Campbell Collaboration²⁾, die nach dem Vorbild der Cochrane Collaboration systematische Reviews in den Sozial- und Verhaltenswissenschaften erstellt. Ebenso sind Bestrebungen in den angelsächsischen und nordeuropäischen Ländern zu beobachten, unter dem Stichwort „Evidence Based Education“ methodische Standards für Bildungsforschung einzuführen. Sowohl für die Biometrie als auch für die Medizinische Informatik kann die interventionelle Bildungsforschung interessante Ansätze darstellen, sowohl für wissenschaftliches Arbeiten, als auch die Umsetzung in praktische Ausbildungsmodule im Interesse der Studierenden.

Literatur

- [1] Auhuber TC, Schaefer HE, Schulz S, Klar R: Computer in der Medizinischen Ausbildung – Kontrollierte Evaluation eines computerbasierten Atlas der Histopathologie. *Medizinische Ausbildung*, 2000; 17: 5–11.
- [2] Boruch RF, de Moya D, and Synder B. in: Mosteller F, Boruch R (Hrsg.). *Evidence Matters Randomized Trials in Education Research*. Brookings Institution Press 2002.
- [3] Boruch RF. The Importance of Randomized Field Trials. *Crime and Delinquency* 2000; 46(2): 156–180.
- [4] Burtless G. Randomized Field Trials for Policy Evaluation: Why Not in Education? In: Frederick Mosteller and Robert Boruch (Hrsg.) *Evidence Matters Randomized Trials in Education Research*. Brookings Institution Press 2002.
- [5] Cook, TD. „Sciencephobia: Why education researchers reject randomized experiments. *Education Next* 2001; 1(3): 62–68.

²⁾ <http://www.campbellcollaboration.org/>

- [6] Senn S. Cross-over Trials in Clinical Research, 2nd Edition, WileyEurope, 2002.
[7] Schulmeister R. Grundlagen Hypermedialer Lernsysteme. 2. Aufl. R. Oldenbourg München 1997.

Abstraktnummer 35/2

Verwendung von computergestützten Fällen als Teil eines OSCE

Körner BH, Fischer M, Holzer M, Schewe S

Medizinische Klinik der Ludwig-Maximilians-Universität München, AG Instruct, München

Fallorientierte Lernsysteme können zur Messung der Fähigkeit Probleme im kognitiven Bereich zu lösen verwendet werden. Wir haben in einer Studie mit 223 Studenten einen Computergestützten Fall mit der Papierversion als eine Station eines OSCE verglichen. Dies ist das erste Mal das unser Web basierten Fallorientieren Lernsystem mit den Namen CASUS innerhalb einer Prüfung verwendet wird.

Alle Studierenden des 4 klinischen Semesters nehmen verpflichtend an einem vier Wochen langen Problem Orientierten Blockkurs teil. Zu den Kursen, die aus Tutorials, Praktischen Übungen, Vorlesungen, „Bedside“ Elementen und Praktika zu Spezialmethoden bestehen, wurden von uns 10 Fälle zum freiwilligen Selbststudium angeboten. Am Ende des Kurses steht ein OSCE mit 4 Stationen, im WS 2002/03 wurden 41 Studenten zufällig ausgewählt, und diesen der Fall in der Prüfung mit dem Computer präsentiert, allen anderen Studenten wurde der Fall, wie in den letzten Semestern, in Papierform vorgelegt.

Die Motivation der Studenten die kursbegleitenden Fälle zu bearbeiten war gut, 56% der Studenten (126 von 223) bearbeiteten, um sich auf die Prüfung vorzubereiten, alle (10) Fälle, nur 12% (27) der teilnehmenden Studenten bearbeiteten keinen der Fälle. Die Studenten die Fälle zur Prüfungsvorbereitung bearbeiteten schnitten in der entsprechenden OCSE Station signifikant besser ab. Entgegen den primären Erwartungen schnitten Studenten die den Fall am Computer bearbeiteten, signifikant schlechter ab, als die Studenten die denselben Fall als Papierversion bearbeiteten. Die Arbeit der Tutoren die Prüfung zu managen und zu bewerten wurde durch die gute Lesbarkeit und eine Vorverarbeitung wesentlich unterstützt.

Für das signifikant schlechtere Abschneiden der Studenten die, die Fragen am Computer bearbeiteten gibt es verschiedene Erklärungsmöglichkeiten. Durch eine Nachauswertung der Prüfungsbögen, konnten einige Studenten identifiziert werden, den trotz der Vielfältigen Angebote noch elementare Computerkenntnisse fehlen. Dies deckt sich mit den Beobachtungen von unseren Observern, die die Studenten während der Prüfung beobachteten. Ein weiterer Einflussfaktor könnte der zur Verfügung stehende Platz gewesen sein. Unserer Meinung nach können Computergestützte Fälle für Prüfungen verwendet werden, aber immer noch gibt es einige wenige Studenten, denen grundlegende Computerkenntnisse fehlen.

Literatur

Bloch et al: IAWF Kompetent Prüfen, Bern Wien 1999.

Fischer MR, Schauer S, Gräsel C, et al.: Modellversuch CASUS. Ein computergestütztes Autorensystem für die Problemorientierte Lehre in der Medizin. *Zeitschrift für Ärztliche Fortbildung*. 1996; 90: 385–389.