



Anleitungen zum Formular für eine kritische Besprechung quantitativer Studien

© Law, M., Stewart, D., Pollock, N., Letts, L., Bosch, J. und Westmorland, M., 1998

Einleitung

- Diese Anleitungen dienen als Begleitinformation zu dem Formular für eine kritische Besprechung quantitativer Studien, welches von der 'Occupational Therapy Evidence-Based Practice Research Group' (Forschungsgruppe für ergotherapeutische Praxis auf wissenschaftlich nachgewiesener Grundlage) (Law et al. 1998) entwickelt wurde. Sie sind in einfacher Sprache verfasst, die sowohl von Forschern als auch von Therapeuten und Studenten verstanden werden kann.
- An geeigneten Stellen werden Beispiele und Begründungen für die Anleitungen/Vorschläge gegeben, damit der Vorgang des kritischen Besprechens dem Leser verständlicher wird.
- Für jede Komponente gibt es Erläuterungen zu den Fragen (linke Spalte des Formulars) und zu den Anweisungen/Fragen in der Kommentarspalte.

Komponenten einer kritischen Besprechung

- Quellenangabe: Geben Sie den vollen Titel, alle Autoren (Nachname, Initialen), den vollen Namen der Zeitschrift, das Erscheinungsjahr, den Band und die Seiten an.
- So ist für andere Personen der Artikel leicht zu finden.

Zweck der Studie

- Wurde der Zweck klar angegeben? – Der Zweck wird gewöhnlich kurz in der Zusammenfassung des Artikels angegeben und dann ausführlicher in der Einleitung. Er kann als Forschungsfrage oder Hypothese formuliert sein.
- Eine klare Formulierung hilft Ihnen, zu entscheiden, ob das Thema wichtig, relevant und für Sie von Interesse ist. Überlegen Sie, bevor Sie fortfahren, wie sich die Studie für die ergotherapeutische Praxis und/oder Ihre eigene Situation nutzen lässt. Ist sie nicht nützlich oder anwendbar, gehen Sie zum nächsten Artikel über.

Literatur

- Wurde die relevante Hintergrund-Literatur gesichtet? – Zu einem Artikel, der ein Forschungsprojekt beschreibt, sollte als gewisser Hintergrund der Studie eine Übersicht über die entsprechende Literatur gehören. Sie sollte eine Synthese der relevanten Informationen – etwa früherer Arbeiten/Forschungen – und eine Diskussion der klinischen Bedeutung des Themas bieten.



- Sie stellt Lücken im heutigen Wissen und der heutigen Forschung bezüglich des fraglichen Themas fest und rechtfertigt so die Notwendigkeit, über die Studie zu berichten.

Design

- Es gibt viele verschiedene Arten von Forschungs-Designs. Nachfolgend werden die üblichsten Arten von Designs in der Rehabilitations-Forschung aufgeführt.
- Die wesentlichen Merkmale der verschiedenen Arten von Studiendesigns werden skizziert, damit Sie leichter bestimmen können, welches Design bei der Studie, die Sie besprechen, verwendet wurde.
- Einige der Vor- und Nachteile der verschiedenen Arten von Designs werden hier skizziert, damit der Leser die Angemessenheit des Designs für die jeweilige publizierte Studie leichter beurteilen kann.
- Verschiedene Autoren verwenden unterschiedliche Begriffe, und dies kann verwirrend sein. Alternative Bezeichnungen werden hier nach Möglichkeit angegeben.
- Zur Beurteilung der Angemessenheit gewählter Methoden/Designs können zahlreiche Aspekte betrachtet werden. Einige der wichtigsten werden im Kommentar-Abschnitt aufgeführt und nachstehend diskutiert. Die verschiedenen Designs werden diagrammatisch dargestellt und an Beispielen aus dem Gebiet der Untersuchung der Wirksamkeit von Aktivitätsprogrammen für Senioren mit Demenz illustriert.
- Die meisten Studien haben gewisse Probleme mit systematischen Fehlern, welche sich verzerrend auf Design, Durchführung oder Interpretation der Forschung auswirken können. Die häufigsten systematischen Fehler werden an späterer Stelle beschrieben.

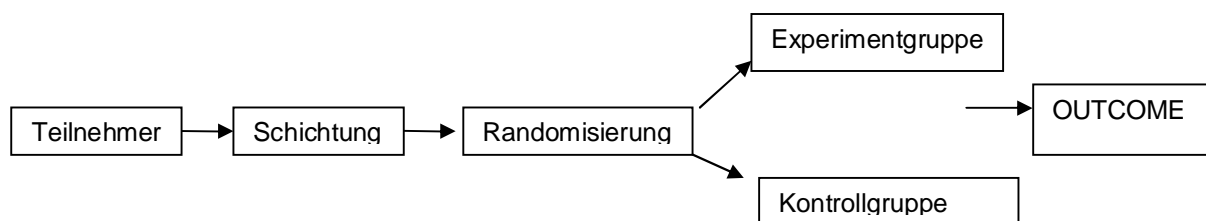
Arten von Designs

1 Randomisierte kontrollierte Studie (RCT)

- Randomisierte kontrollierte Studien / randomisierte klinische Versuchsreihen (RCTs) werden auch als experimentelle Studien oder Studien vom Typ 1 bezeichnet. RCTs umfassen auch andere verschiedene Methoden, etwa Cross-over-Designs.
- Das wesentliche Merkmal einer RCT besteht darin, dass eine Reihe von Klienten/Studiensubjekten gefunden wird, die dann nach dem Zufallsprinzip zwei oder mehr Behandlungs-"Gruppen" zugewiesen werden. Die Klienten in der einen Gruppe werden in der interessierenden Weise behandelt (oft eine neue Behandlungsform), und die andere Gruppe ist die "Kontroll"-Gruppe, die gewöhnlich gar nicht oder mit Standard-Verfahren behandelt wird. Zufällige Zuweisung zu verschiedenen Behandlungsgruppen ermöglicht den Vergleich der Klientengruppen hinsichtlich der Ergebnisse der Behandlung, denn dadurch nimmt die Wahrscheinlichkeit sehr stark zu, dass es in jeder Gruppe ähnliche Klienten gibt. Die Möglichkeit,

dass irgendein anderer Faktor (eine sogenannte "Störvariable" oder ein "Störaspekt" (confounding variable/issue)) die Resultate beeinflusst, nimmt also sehr stark ab.

- Der Hauptnachteil von RCTs sind die damit verbundenen Kosten. Außerdem ist es in manchen Fällen ethisch nicht zu verantworten, "Kontroll"-Gruppen von Klienten zu haben, die nicht behandelt werden. Wäre beispielsweise die Wirksamkeit eines multidisziplinären Programms für stationäre frisch operierte Patienten mit chronischen Kreuzschmerzen zu untersuchen, so wäre es unter Umständen unethisch, einigen Patienten eine Behandlung vorzuenthalten, um eine "Kontroll"-Gruppe zu haben.
- RCTs werden oft zum Testen der Wirksamkeit einer Behandlung gewählt oder zum Vergleich verschiedener Behandlungsformen.



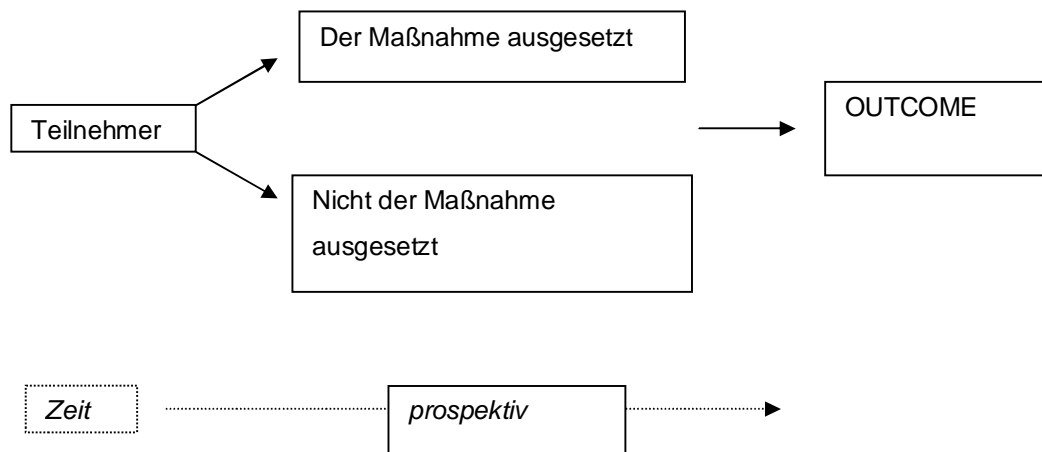
Beispiel: Die Wirkungen zweier verschiedener ergotherapeutischer Maßnahmen – funktioneller Rehabilitation und Reaktivierung – wurden mit einer randomisierten kontrollierten Versuchsreihe bewertet. 44 Patienten in einem Langzeit-Betreuungszentrum wurden nach dem Zufallsprinzip einer der beiden Behandlungsformen zugeteilt. Die Ausgänge wurden zu drei verschiedenen Zeitpunkten gemessen, mit verschiedenen psychometrischen Tests (Bach et al., 1995).

2 Kohortendesign

- Eine Kohorte ist eine Gruppe von Menschen (Klienten), die einer ähnlichen Situation ausgesetzt (exponiert) sind, z.B. Menschen mit ähnlichem Behandlungsprogramm oder ähnlicher Diagnose/Krankheit. Zu dem jeweiligen Forschungsthema / der jeweiligen Frage wird eine entsprechende Gruppe von Klienten identifiziert und über einen bestimmten Zeitraum hinweg begleitet/beobachtet, so dass gesehen werden kann, was passiert.
- Kohorten-Designs sind "prospektiv", das heißt, immer in die Zukunft gerichtet. Es geht um die Zeit vom Moment der Identifizierung der Klienten an. Manchmal werden solche Studien auch als prospektive Studien bezeichnet.
- Bei Kohortenstudien gibt es oft ebenfalls eine Vergleichs- ("Kontroll"-) -Gruppe von Klienten/Personen, die der fraglichen Situation nicht ausgesetzt waren (die z.B. überhaupt nicht

behandelt wurden). Einer der Hauptunterschiede zwischen einer RCT und einer Kohortenstudie besteht darin, dass bei einer Kohortenstudie die Zuteilung der Personen (Klienten) zu Behandlungs- oder Kontrollgruppe nicht der Kontrolle des Forschers unterliegt – der Forscher muss mit der Gruppe von Personen arbeiten, die als "exponiert" identifiziert wurden, und dazu eine Gruppe anderer Personen finden, die hinsichtlich Alter, Geschlecht und weiterer wichtiger Faktoren ähnlich sind.

- Da es schwierig ist, zu wissen, ob die Gruppen hinsichtlich aller wichtigen (bzw. Stör-) Faktoren ähnlich sind, können die Autoren nicht mit Bestimmtheit sagen, dass wirklich die fragliche Behandlung (Exposition) für das Ergebnis verantwortlich ist.
- Ein Vorteil von Kohortenstudien liegt darin, dass sie oft weniger kostspielig und zeitaufwendig sind als RCTs.

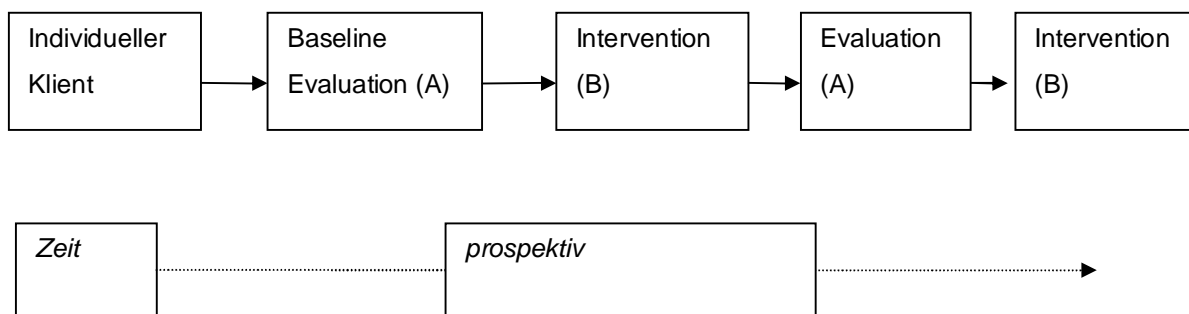


Beispiel: Zur Bewertung eines Programms zur mentalen Anregung wurden in einer Kohortenstudie bei 30 Patienten über einen Zeitraum von 2 Monaten Veränderungen des mentalen Status gemessen. Die ersten 15 Patienten, die in ein Tagesbetreuungs-Zentrum aufgenommen wurden, wurden behandelt und bildeten die Gruppe der "Exponierten". Die übrigen 15 aufgenommenen Personen wurden nicht sofort behandelt und dienten als "Kontroll"-Gruppe (Koh et al., 1994).

3 Einzelfalldesign

- Bei Einzelpersonen-/Einzelfall-Forschung werden ein Klient oder eine Reihe von Klienten über einen bestimmten Zeitraum beobachtet oder bezüglich interessierender Ergebnisse beurteilt.

- Es gibt für Einzelfalldesigns verschiedene Arten von Methoden und verschiedene Begriffe, etwa "N=1"-Studie, "Vorher-Nachher-Studie beim selben Subjekt" oder Einzelfall-"Serie" (bei der mehr als eine Person / ein Klient beteiligt sind).
- Das grundlegende Merkmal jedes Einzelpersonen-Designs ist die Beurteilung von Klienten hinsichtlich der interessierenden Ausgänge sowohl vor der Maßnahme (Anfangswert) als auch nachher. Bei dieser Versuchsanordnung kann eine Person sich selbst als "Kontrolle" dienen. Daraus lässt sich jedoch schwerlich schließen, dass allein die Behandlung für irgendwelche Unterschiede verantwortlich ist, da sich auch andere Faktoren im Laufe der Zeit ändern können, etwa der Schweregrad einer Erkrankung.
- Dieses Studiendesign ist nützlich, wenn eine Diagnose oder Behandlungsform, die man evaluieren will, nur bei einigen wenigen Klienten anzutreffen ist. Eine derartige Studie lässt sich leicht mit mehr als einem Klienten wiederholen. Wegen seines flexiblen Ansatzes eignet sich das Design besonders für Forschung im klinischen Rahmen.

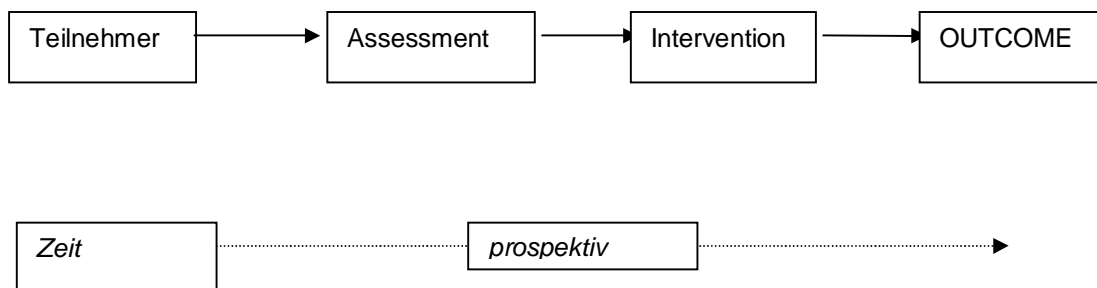


Beispiel: Zur Untersuchung der Wirkungen von Veränderungen des Umfelds während ergotherapeutischer Maßnahmen in einer psychiatrischen Station wurden gemäß einem Einzelfalldesign Verhaltensänderungen bei 10 einzelnen Patienten beobachtet. Beobachtungen des Verhaltens jedes Patienten wurden vor, während und nach der Maßnahme gemacht (Burton, 1980).

4 Vorher-Nachher-Design

- Ein Vorher-Nachher-Design wird gewöhnlich zur Beurteilung der Behandlungswirkungen bei einer Gruppe von Klienten eingesetzt (wenn es auch, wie bereits erwähnt, ebenfalls als Methode zur Untersuchung von Einzelfällen/Einzelpersonen verwendet wird).

- Der Untersucher sammelt bei einer Gruppe von Klienten Informationen über deren anfänglichen Status hinsichtlich der interessierenden Ergebnisse und sammelt dann nach der Behandlung erneut Informationen über diese Ergebnisse.
- Dieses Design ist nützlich, wenn man keinem Klienten die Behandlung vorenthalten will. Ohne "Kontroll"-Gruppe ist es allerdings unmöglich, zu beurteilen, ob allein die Behandlung für irgendwelche Veränderungen der Ausgänge verantwortlich war. Veränderungen könnten auch auf andere Faktoren zurückzuführen sein, etwa auf das Fortschreiten der Krankheit, die Medikation, den Lebensstil oder Veränderungen des Umfelds.



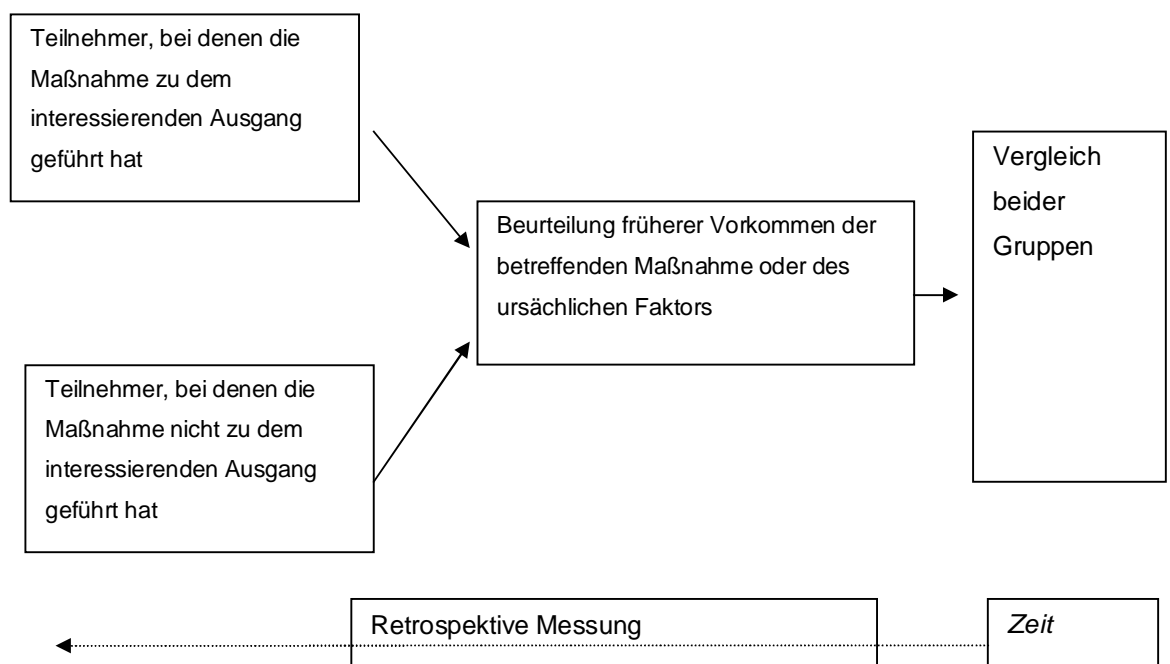
Beispiel: Die Belastung durch Betreuung eines älteren Familienmitgliedes mit Demenz wurde gemäß einem Vorher-Nachher-Design beurteilt. Die Belastung des Betreuers wurde bei 15 Personen gemessen, bevor und nachdem das jeweilige Familienmitglied in eine Tagesbetreuungs-Einrichtung aufgenommen worden war (Graham, 1989).

5 Fall-Kontroll-Design

- Fall-Kontroll-Studien erforschen, worin sich eine Gruppe von Personen unterscheidet. Sie werden auch als Fall-Vergleichs-Studien oder retrospektive Studien bezeichnet. Der Begriff 'retrospektiv' wird verwendet, um zu bezeichnen, dass diese Methoden etwas im Rückblick anschauen, nachdem es geschehen ist. Das wesentliche Merkmal einer Fall-Kontroll-Studie ist ihr Blick zurück.
- Es werden eine Reihe von Klienten/Personen identifiziert, auf die ein bestimmtes Merkmal oder eine Situation zutrifft, etwa eine spezifische Diagnose oder die Teilnahme an einer Behandlung.

Bezüglich des betreffenden Merkmals oder der Situation wird mit einer "Kontroll"-Gruppe von Personen verglichen, die in Alter, Geschlecht und Hintergrund ähnlich sind, aber das Merkmal nicht aufweisen bzw. nicht in der gleichen Situation sind. Zweck ist es, die Unterschiede zwischen diesen Gruppen zu bestimmen.

- Dieses Verfahren ist eine relativ kostengünstige Art, eine Frage zu erforschen. Es weist aber viele potentielle Probleme (Mängel) auf, aufgrund derer sich nur schwerlich schließen lässt, welche Faktoren für den Ausgang verantwortlich sind.

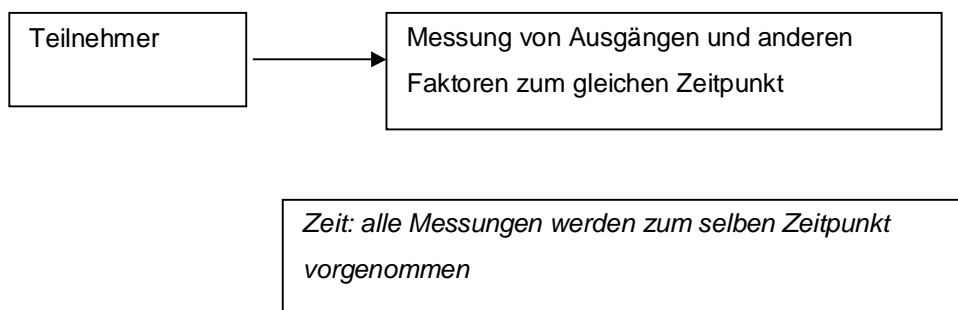


Beispiel: Falls ein Ergotherapeut herausfinden wollte, warum manche Klienten in einer Tagesbetreuungs-Einrichtung an einem täglichen (freiwilligen) Aktivitätsprogramm regelmäßig teilgenommen haben und andere nicht, so könnte er die Unterschiede zwischen den Klientengruppen hinsichtlich Alter, Geschlecht, Interessen, Hintergrund und aktueller Lebenssituation mit einem Fall-Kontroll-Design untersuchen.

6 Querschnitts-Design

- Bei diesem Design gibt es nur eine Gruppe von Personen, und die Bewertung der gesamten Gruppe erfolgt zum gleichen Zeitpunkt.

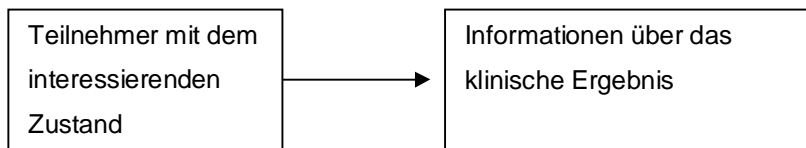
- Dieses Design wird oft verwendet, um zu erforschen, welche Faktoren sich auf einen bestimmten resultierenden Zustand bei einer Gruppe von Personen ausgewirkt haben könnten. Ein solches Design ist nützlich, wenn man über das betreffende Problem / den Zustand relativ wenig weiß.
- Umfragen, Fragebogen und Interviews sind übliche Methoden bei Querschnitts-Studien. Sie kosten relativ wenig und sind leicht durchzuführen, da die Bewertung zu einem einzigen Zeitpunkt stattfindet.
- Es ist allerdings unmöglich, zu wissen, ob alle Faktoren in die Bewertung einbezogen wurden. Daher lassen sich aus den Resultaten schwerlich Schlussfolgerungen über Ursache-Wirkungs-Zusammenhänge ziehen, die über den Rahmen der untersuchten Gruppe hinausgehen.



Beispiel: Zur Bewertung eines neuen Aktivitätsprogramms für Senioren mit Demenz können bei der Entlassung Klienten, die an dem Programm teilgenommen haben, sowie deren Familien mit einer Umfrage oder mit Interviews nach der Auswirkung des Programms auf die Lebensqualität, die Teilnahme an Aktivitäten und den Grad der Zufriedenheit befragt werden.

7 Fallstudien-Design

- Eine Fallstudie wird durchgeführt, um deskriptive Informationen (Daten) über die Beziehung zwischen einer bestimmten Behandlung (Exposition) und einem interessierenden Ergebnis zu gewinnen. Eine solche Studie wird auch als deskriptive Studie bezeichnet, weil dies ihr primärer Zweck ist. Eine Kontrollgruppe gibt es nicht.
- Eine Fallstudie wird oft eingesetzt, um ein neues Thema oder eine neue Behandlung zu erforschen, wenn man darüber noch wenig weiß. Die Ergebnisse lassen sich jedoch nur als Beschreibung einer spezifischen Situation ansehen. Eine solche Studie kann Informationen beschaffen, die für weitere Studien zu dem interessierenden Thema nützlich sind.

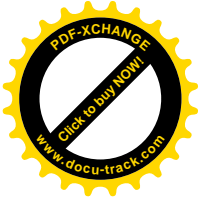
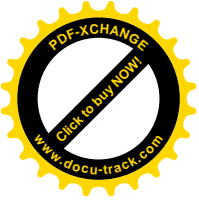


Beispiel: 12 Patienten auf einer geriatrischen Langzeit-Station wurden über einen bestimmten Zeitraum beobachtet, um die Wirksamkeit des Angebots von Einzel- und Gruppenaktivitäten auf der Station zu beurteilen. Der Grad ihrer Beteiligung wurde in 10-Minuten-Intervallen aufgezeichnet, um mögliche Unterschiede zwischen keinerlei Maßnahmen, Einzelaktivitäten und Gruppenaktivitäten festzustellen (McCormack & Whitehead, 1981).

Angemessenheit des Studiendesigns

Nachfolgend einige der wichtigen Punkte, die bedacht werden müssen, um zu bestimmen, ob ein Studiendesign das geeignetste ist:

- Kenntnis des Themas / der Frage: Wenn über ein Thema wenig bekannt ist, eignet sich eine eher exploratorische Methode, z.B. eine Fallstudie oder ein Querschnitts-Design. Mit zunehmendem Wissen werden Studiendesigns strenger – man kennt die meisten Variablen, die den Ausgang beeinflussen könnten, und kann sie kontrollieren. Das strengste Design ist die RCT.
- Ergebnisse/outcome: Lässt sich der untersuchte outcome von Maßnahmen leicht quantifizieren und gibt es dazu gut entwickelte, standardisierte Messinstrumente, so eignet sich ein strengeres Design (z.B. eine RCT). Versteht man Ergebnisse noch nicht vollständig (wie etwa Lebensqualität), dann eignet sich ein Design, das verschiedene Faktoren erforscht, welche für diese Ausgänge eine Rolle spielen könnten, beispielsweise ein Fall-Kontroll-Design.
- Ethische Fragen: Wenn es keine ethischen Gründe gibt, die gegen das Unterlassen einer Behandlung sprechen, ist es angemessen, ein Forschungs-Design zu verwenden, bei dem Kontrollgruppen von Personen nicht behandelt werden.
- Studienzweck/Studienfrage: Manche Designs eignen sich gut, um die Wirksamkeit einer Behandlung zu untersuchen, etwa RCTs, Vorher-Nachher-Designs und Einzelfallstudien. Andere Designs (z.B. Fall-Kontroll-Studien und Querschnittsstudien) eignen sich eher bei Untersuchungen, deren Zweck es ist, mehr über ein Thema herauszufinden, oder für Pilotstudien, die klären sollen, ob eine bestimmte Behandlung oder Forschung weiterhin gerechtfertigt ist.

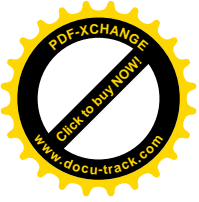


Systematische Fehler

- Viele verschiedene Arten systematischer Fehler werden in der Forschungsliteratur beschrieben. Die häufigsten (auf die hin Artikel überprüft werden sollten) werden nachfolgend, untergliedert in 3 Hauptgebiete, beschrieben:
 - systematische Fehler bei der Stichprobe / Auswahl der Subjekte, infolge derer die Subjekte in der Stichprobe vielleicht nicht repräsentativ sind für die Population, an der man interessiert ist
 - systematische Fehler beim Messen/Ermitteln, die damit zusammenhängen, wie ein fraglicher Ausgang gemessen wurde
 - systematische Fehler bei den Maßnahmen / der Durchführung, die damit zusammenhängen, wie die Behandlung selbst durchgeführt wurde
- Für detailliertere Informationen über systematische Fehler sei der Leser auf die Literaturangaben verwiesen.
- Ein systematischer Fehler beeinflusst die Ergebnisse einer Studie in einer Richtung – entweder "begünstigt" er die Behandlungsgruppe oder die Kontrollgruppe. Es ist wichtig, sich bewusst zu werden, in welche Richtung ein systematischer Fehler die Ergebnisse beeinflussen kann.

1 Systematische Fehler bei Stichprobe/Auswahl

- a hinsichtlich Freiwilliger oder Überweisungen
 - Personen, die freiwillig an einer Studie teilnehmen oder durch jemanden an die Studie überwiesen werden, unterscheiden sich oft von nicht freiwilligen / nicht überwiesenen Personen.
 - Dieser systematische Fehler begünstigt gewöhnlich (wenn auch nicht immer) die Behandlungsgruppe, weil Freiwillige tendenziell motivierter und mehr um ihre Gesundheit besorgt sind.
- b hinsichtlich der Jahreszeit
 - Wenn alle Personen zur gleichen Zeit rekrutiert und also auch behandelt und beurteilt werden, können die Ergebnisse durch den Zeitpunkt der Personenauswahl und der Maßnahmen beeinflusst sein. Zum Beispiel sind Senioren im Sommer tendenziell gesünder als im Winter; daher können die Ergebnisse positiver ausfallen, wenn die Studie nur im Sommer durchgeführt wird.
 - Dieser systematische Fehler kann sich, je nach Jahreszeit, in beide Richtungen auswirken.
- c hinsichtlich der Aufmerksamkeit
 - Personen, die im Rahmen einer Studie beurteilt werden, kennen gewöhnlich den Zweck der Studie und geben infolge der ihnen gewidmeten Aufmerksamkeit günstigere Antworten oder zeigen bessere Leistungen als Personen, welche die Absicht der Studie



nicht kennen. Wegen dieses systematischen Fehlers gibt es in manchen Studien eine "Aufmerksamkeits-Kontrollgruppe" – die Personen in der Kontrollgruppe erfahren dort ebenso viel Aufmerksamkeit wie jene in der Behandlungsgruppe, obwohl die Behandlung nicht die gleiche ist.

2 Systematische Fehler beim Messen/Ermitteln

- a hinsichtlich der Zahl der verwendeten Maße für den Ausgang
- Wird nur ein Maß für den Ausgang verwendet, kann es einen systematischen Fehler geben bei der Art, wie jenes Maß den Ausgang beurteilt. So beachtet zum Beispiel ein bestimmtes ADL-Maß das Anziehen, Essen und den Gang zur Toilette, bezieht aber Körperpflege und Pflege der Kleidung oder Zubereiten von Mahlzeiten nicht ein.
 - Dieser systematische Fehler kann die Ergebnisse in beide Richtungen beeinflussen; er kann z.B. die Kontrollgruppe begünstigen, wenn wichtige Elemente des outcomes, welche auf die Behandlung angesprochen hätten, nicht einbezogen werden.
 - Zu einem systematischen Fehler kann es auch kommen, wenn es angesichts der Stichprobengröße zu viele outcome Messungen gibt. Dieser Punkt betrifft die statistische Auswertung. Gewöhnlich wird dadurch die Kontrollgruppe begünstigt, denn eine große Zahl statistischer Berechnungen verringert die Aussicht, einen signifikanten Unterschied zwischen Behandlungs- und Kontrollgruppe zu finden.
- b wegen des Fehlens "maskierter" oder "unabhängiger" Bewertung
- Wissen die Beurteiler, welcher Gruppe eine Person zugeteilt war oder wie sie behandelt wurde, so können sie unter Umständen die Ergebnisse beeinflussen, indem sie die Person oder Gruppe von Personen mehr oder weniger günstiger beurteilen. Gewöhnlich wird dadurch die Behandlungsgruppe begünstigt. Daran sollte gedacht werden, wenn die Beurteiler zum Forschungs- oder Behandlungsteam gehören.
- c hinsichtlich der Erinnerung
- Dies kann ein Problem sein, wenn Ergebnisse gemessen werden anhand von eigenen Angaben der Studiensubjekte, Umfragen oder Interviews, bei denen sich die Befragten an frühere Ereignisse erinnern müssen. Oft erinnern sich Menschen an Zuversicht vermittelnde oder positive Dinge mehr als an negative, und dadurch können die Ergebnisse einer Studie zugunsten jener Personen ausfallen, die über einen Punkt befragt oder die behandelt wurden.

3 Systematische Fehler bei Maßnahmen/Durchführung

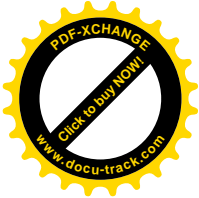
- a Kontaminierung



- Kontaminierung findet statt, wenn Mitglieder der Kontrollgruppe versehentlich behandelt werden. Dadurch kann sich der Unterschied zwischen den beiden Gruppen verringern. Dies begünstigt die Kontrollgruppe.
- b gleichzeitige weitere Maßnahmen (Ko-Intervention)
 - Wenn Klienten gleichzeitig mit der untersuchten Behandlung auch noch in anderer Form behandelt werden, kann dies die Ergebnisse in beiden Richtungen beeinflussen. Beispielsweise kann die Einnahme von Medikamenten während der Behandlung oder Nicht-Behandlung die Ergebnisse zugunsten beider Gruppen beeinflussen. Der Leser muss abwägen, ob die andere oder zusätzliche Behandlung einen positiven oder negativen Einfluss auf die Ergebnisse gehabt haben könnte.
- c zeitliche Aspekte der Maßnahmen
 - Im Zusammenhang mit zeitlichen Aspekten der Maßnahmen können verschiedene Punkte systematische Fehler hervorrufen.
 - Wenn Kinder über längere Zeit behandelt werden, könnte schon allein der Reifeprozess bei den beobachteten Verbesserungen eine Rolle spielen.
 - Dauert die Behandlung nur sehr kurze Zeit, so hat sich vielleicht eine erkennbare Wirkung bezüglich der interessierenden Ergebnisse noch nicht zeigen können. Dies würde die Kontrollgruppe begünstigen.
- d Ort der Behandlung
 - Der Ort, wo eine Behandlung durchgeführt wird, kann die Ergebnisse beeinflussen. Wenn beispielsweise eine Person zu Hause behandelt wird, so kann dadurch vielleicht ihre Zufriedenheit größer sein; dies begünstigt die Behandlungsgruppe. Der Ort der Behandlung sollte bei allen Gruppen übereinstimmen.
- e verschiedene Therapeuten
 - Wenn die untersuchte Behandlung in den verschiedenen Klientengruppen von verschiedenen Therapeuten ausgeführt wird, so kann dies die Ergebnisse in einer Richtung beeinflussen. Beispielsweise könnte der eine Therapeut motivierender oder positiver sein als ein anderer, und daher könnte die Gruppe, mit der er arbeitet, günstigere Resultate aufweisen. Die Beteiligung der Therapeuten sollte für alle Behandlungsgruppen gleich und konsistent sein.

Stichprobe

- N = ? Die Anzahl der an der Studie beteiligten Personen/Klienten sollte klar sein.
- Wurde die Stichprobe detailliert beschrieben? Die Beschreibung der Stichprobe sollte so genau sein, dass man ein klares Bild davon bekommt, wer beteiligt war.



- Merkmale, die hinsichtlich des interessierenden Themas wichtig sind, sollten angegeben sein, damit Sie feststellen können, ob die Studienpopulation Ihrer eigenen ähnelt und ob systematische Fehler so gering wie möglich gehalten wurden. Wichtig sind zum Beispiel folgende Aspekte:
 - Aus welchen Personen setzt sich die Stichprobe zusammen – eignen sich diese Personen bezüglich der Studienfrage, und wurde ihr Alter, Geschlecht, die Dauer ihrer Behinderung/Krankheit und ihr funktioneller Status (falls relevant) angegeben?
 - Wie viele Personen waren beteiligt, und – falls es mehrere Gruppen gab – hatten die Gruppen eine relativ gleiche Größe?
 - Wie wurde die Stichprobe zusammengestellt – aus freiwilligen, aus überwiesenen Personen? Sind Einschluss- und Ausschlusskriterien beschrieben?
 - Bei mehr als einer Gruppe: Waren sich die Gruppen bezüglich wichtiger (Stör-) Faktoren ähnlich?
- Wurde die Stichprobengröße begründet? Die Autoren sollten angeben, wie sie zu der Stichprobengröße kamen, sie sollten begründen, warum diese Teilnehmerzahl gewählt wurde. Oft wird die Stichprobengröße mit der für die Studie verfügbaren Population begründet. Manche Autoren geben statistische Gründe an, aber das ist selten.
- Das Ethik-Verfahren sollte beschrieben sein, auch wenn dies oft unterlassen wird. Allermindestens sollten die Autoren angeben, ob zu Beginn der Studie die wohlinformierte Zustimmung der Teilnehmer eingeholt wurde.

Ergebnis / outcomes

- outcomes sind die den Forscher interessierenden Variablen oder Punkte – sie repräsentieren das Produkt oder die Resultate der Behandlung oder Exposition.
- outcomes müssen klar beschrieben sein, damit Sie beurteilen können, ob sie für Ihre Situation relevant und nützlich sind. Außerdem sollten die Methoden (das Wie) der Messungen von outcomes hinlänglich beschrieben sein, so dass darauf vertraut werden kann, dass sie in objektiver und unvoreingenommener Weise durchgeführt wurden.
- Stellen Sie die Häufigkeit der Messungen von outcomes fest. Es ist wichtig, zu beachten, ob outcomes vor und nach Behandlung gemessen wurden und ob kurzfristige und/oder langfristige Wirkungen bedacht wurden.
- Überprüfen Sie die Maße für outcomes darauf, inwiefern sie für die ergotherapeutische Praxis relevant sind, d.h., ob sie sich auf die Bereiche Durchführungsqualität einer Tätigkeit (occupational performance), Komponenten der Performanz (performance components) und/oder Komponenten des Umfelds (environmental components) beziehen.
- Listen Sie die verwendeten Maße und alle wichtigen Informationen über sie auf, um zukünftig darauf Bezug nehmen zu können. Überlegen Sie, ob es sich um wohlbekannte Maße handelt

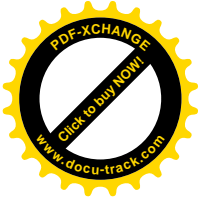


oder solche, die die Forscher spezifisch für die betreffende Studie entwickelt haben. In diesem Fall könnte es schwieriger sein, die Studie in einer späteren Situation zu wiederholen.

- Die Autoren sollten angeben, ob die verwendeten Maße für Ausgänge solide (woherprobte) psychometrische Eigenschaften haben – vor allem Reliabilität und Validität. Dies gewährleistet die Verlässlichkeit der betreffenden Messungen der Ausgänge.
- Waren die Maße für outcomes zuverlässig (reliabel)? – Reliabilität gibt an, ob ein Maß in unterschiedlichen Situationen dieselbe Information liefert. Die beiden üblichsten Formen von Reliabilität sind: Test-Retest-Reliabilität (derselbe Beobachter ermittelt an zwei zeitlich kurz auseinanderliegenden Gelegenheiten dieselbe Information) und Interbeurteiler-Reliabilität (verschiedene Beobachter ermitteln zur selben Zeit dieselbe Information).
- Waren die Maße für outcomes gültig (valide)? – Hier wird danach gefragt, ob ein Maß wirklich das beurteilt, was es messen soll. Achten Sie darauf, ob das Maß alle relevanten Konzepte und Elemente des fraglichen outcomes einbezieht (Inhalts-Validität) und ob die Autoren angeben, das Maß sei in Bezug zu andern Maßen getestet, so dass sein Verhältnis zu jenen klar ist (Kriteriumsvalidität). Ein "valides" ADL-Maß zum Beispiel erstreckt sich auf alle relevanten Elemente der Selbstversorgung (self care) und wurde anhand anderer Maße für Aktivitäten des täglichen Lebens bzw. für Fähigkeiten zur Selbstversorgung getestet, so dass es nachweislich zu ihnen im erwarteten Verhältnis steht.

Maßnahmen

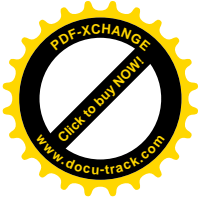
- Wurden die Maßnahmen detailliert beschrieben? – Die Informationen über die Maßnahmen sollten ausreichen, damit Sie diese nachvollziehen können.
- Bedenken Sie bei der Begutachtung der Maßnahmen wichtige Elemente wie die folgenden:
 - Schwerpunkt der Maßnahme – ist er für die ergotherapeutische Praxis und Ihre Situation relevant?
 - Wer hat die Maßnahme durchgeführt – eine einzige Person oder mehrere Personen, war(en) sie ausgebildet?
 - Wie oft wurde behandelt – genügte das Ihrer Meinung nach, um Wirkung zu zeigen? Wurde, wenn es mehrere Gruppen gab, in allen gleich häufig behandelt?
 - Rahmen – wurde zu Hause oder in einer Institution behandelt? War dies, falls es mehr als eine Behandlungsgruppe gab, für die verschiedenen Gruppen gleich?
- Diese Elemente müssen angegeben sein, wenn Sie in der Lage sein sollen, die Behandlung in Ihrer Praxis zu wiederholen.
- Wurde Kontamination, Ko-Intervention vermieden? – Diese beiden Faktoren werden im Abschnitt 'Systematische Fehler' beschrieben. Wurden darüber Angaben gemacht? Wenn nicht, überlegen Sie, welche möglichen Punkte die Resultate der Studie beeinflusst haben könnten.



Was hätte zum Beispiel geschehen können, wenn einige der Klienten in der Kontrollgruppe versehentlich behandelt worden wären (Kontaminierung) oder wenn einige Personen während der Studie Medikamente genommen hätten (Ko-Intervention)? Achten Sie auf alle potentiellen Einflüsse. Wenn nur eine Gruppe untersucht wurde, tragen Sie "entfällt" (N/A = not applicable) in das Formular ein.

Ergebnisse

- Wurde die statistische Signifikanz der Ergebnisse angegeben? – Die meisten Autoren geben zu den Ergebnissen quantitativer Forschung deren statistische Signifikanz an, um nachzuweisen, dass die Resultate Aufmerksamkeit verdienen. Wenn nur Mittelwerte oder Prozentsätze angegeben werden, lässt sich schwer feststellen, ob Veränderungen bei den Ausgängen oder Unterschiede zwischen Personengruppen wichtig oder signifikant sind.
- Bezüglich spezifischer statistischer Methoden sei der Leser auf die Literaturangaben verwiesen.
- Skizzieren Sie in diesem Abschnitt kurz die Ergebnisse, mit Schwerpunkt auf jenen, die statistisch signifikant waren. Waren die Ergebnisse nicht statistisch signifikant, untersuchen Sie die Gründe: War vielleicht die Stichprobe nicht groß genug, um wichtige oder signifikante Unterschiede zu zeigen? Oder wurden, im Verhältnis zu der Zahl der beteiligten Personen, zu viele Maße für den Ausgang verwendet?
- War(en) die Analyse(n) geeignet? – Begründen/Erläutern die Autoren ihre Wahl der Analysemethoden? Scheinen diese Methoden geeignet im Hinblick auf die Studie und deren Ausgänge? Dazu müssen Sie Folgendes beachten:
 - Zweck der Studie – vergleicht sie zwei oder mehrere Maßnahmen, oder untersucht sie die Korrelation zwischen verschiedenen interessierenden Variablen? Für Vergleich bzw. Korrelation werden unterschiedliche statistische Tests verwendet.
 - Ergebnis/outcome– wird nur ein Ergebnis gemessen, um zwei unterschiedliche Behandlungen zu vergleichen, genügt wahrscheinlich ein einfacher statistischer Test wie der t-Test. Bei einer größeren Zahl von Ergebnissen, bei denen unterschiedliche Arten von Variablen eine Rolle spielen, sind hingegen gewöhnlich komplexere statistische Methoden wie Varianzanalyse (ANOVA) erforderlich.
- Wurde die klinische Bedeutung angegeben? – Zahlen genügen oft nicht, um feststellen zu können, ob die Ergebnisse einer Studie klinisch wichtig sind. Die Autoren sollten die Relevanz der Ergebnisse für die klinische Praxis und/oder für das Leben der betroffenen Personen erörtern. Falls signifikante Unterschiede zwischen Behandlungsgruppen gefunden wurden: Sind sie im therapeutischen Zusammenhang aussagekräftig? Falls die Unterschiede nicht statistisch signifikant waren: Gibt es klinisch wichtige oder aussagekräftige Punkte, die für Ihre Praxis in Frage kommen könnten?



Drop-outs / Ausscheiden von Teilnehmern

- Wurden Fälle von Ausscheiden aus der Studie angegeben? – Die Zahl der Personen/Teilnehmer, die aus einer Studie ausscheiden, sollte angegeben sein, denn sie kann die Ergebnisse beeinflussen. Gründe für das Ausscheiden und Angaben darüber, inwiefern Ausgeschiedene bei der Analyse der Ergebnisse berücksichtigt wurden, sollten aufgeführt sein, um die Vertrauenswürdigkeit der Ergebnisse zu erhöhen. Gab es keine Ausgeschiedenen, so betrachten Sie dies als 'angegeben' und vermerken im Kommentarabschnitt 'keine Drop-outs'.

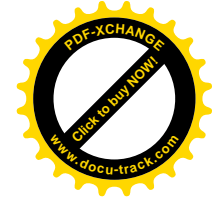
Schlussfolgerungen und klinische Implikationen

- Im Abschnitt 'Diskussion' des Artikels sollten klare Schlussfolgerungen aus den Ergebnissen skizziert sein. Diese sollten im Hinblick auf die Methoden und Ergebnisse der Studie relevant und angemessen sein. Zum Beispiel könnten die Untersucher einer gut konzipierten RCT mit soliden Maßen für den Ausgang darlegen, dass die Ergebnisse den Schluss zulassen, dass Behandlung A für die Studienpopulation wirksamer ist als Behandlung B. Bei anderen Studiendesigns können nicht solche starken Schlüsse gezogen werden, denn wahrscheinlich ist es wegen methodischer Begrenzungen oder systematischer Fehler, etwa dem Fehlen einer Kontrollgruppe oder nicht verlässlichen Maßen, schwierig, zu "beweisen" oder zu schließen, dass allein die fragliche Behandlung die erzielten Ausgänge beeinflusst hat. In diesen Situationen können die Autoren vielleicht nur den Schluss ziehen, dass die Ergebnisse für die beteiligten Klienten einen Unterschied hinsichtlich der spezifischen, in der Studie gemessenen Ausgänge aufzeigen. Die Ergebnisse sind dann vielleicht nicht auf andere Populationen verallgemeinerbar, vielleicht auch nicht auf die Ihre. Daher sollte weitere Forschung empfohlen werden.
- In der Diskussion sollte auch erörtert werden, wie die Ergebnisse die klinische Praxis beeinflussen könnten. Liefern sie nützliche und relevante Informationen über eine Klientenpopulation oder einen interessierenden outcome? Rechtfertigen sie weitere Studien? Überlegen Sie, welche Implikationen die Ergebnisse insgesamt oder Teile der Ergebnisse für Ihre spezifische Praxis bzw. für die Ergotherapie generell haben.

Literaturangaben

Crombie, I.K. (1996). The pocket guide to critical appraisal: A handbook for health care professionals. London: BMJ Publishing Group.

Dept. Of Clinical Epidemiology and Biostatistics, McMaster University Health Sciences Centre (1981). How to read clinical journals: V: To distinguish useful from useless or even harmful therapy. Canadian Medical Association Journal, 124, 1156-1162.



Law, M. (1987). Measurement in occupational therapy: Scientific criteria for evaluation. *Canadian Journal of Occupational Therapy*, 58, 171-179.

Law, M., King, G. & Pollock, N. (1994). *Single Subject Research Design*. Research Report #94- 2. Hamilton, ON: Neurodevelopmental Clinical Research Unit.

Mulrow, C.D., Oxman, A.D. (Eds.). (1996). *Cochrane Collaboration Handbook*. Available in The Cochrane Library [database on disk and CD-ROM]. The Cochrane Collaboration: Issue 2. Oxford: Updated Software; 1998.

Norman, G.R. & Streiner, D.L. (1986). *PDQ Statistics*. Burlington, ON: B.C. Decker Inc.

Sackett, D. L. (1979). Bias in analytic research. *Journal of Chronic Disability*, 32, 51-63.

Sackett, D.L., Haynes, R.B., Guyatt, G.H. & Tugwell, P. (1991). *Clinical Epidemiology. A Basic Science for Clinical Medicine*. Second Edition. Toronto, ON: Little, Brown and Co.

Streiner, D.L., Norman, G.R. & Blum, H.M. (1989). *PDQ Epidemiology*. Toronto, ON: B.C. Decker Inc.

Artikel über Aktivitätsprogramme für Senioren mit Demenz (auf die in Beispielen für Studiendesigns Bezug genommen wurde)

Bach, D., Bach, M., Bohmer, G., Gruhwalk, T. & Grik, B. (1995). Reactivating occupational therapy: a method to improve cognitive performance in geriatric patients. *Age and Aging*, 24, 222-226.

Burton, M. (1980). Evaluation and change in a psychogeriatric ward through direct observation and feedback. *British Journal of Psychiatry*, 137, 566-571.

Graham, R.W. (1989). Adult day care: how families of the dementia patient respond. *Journal of Gerontological Nursing*, 15(3), 27-31, 40-41.

Koh, K., Ray, R., Lee, J., Nair, T., Ho, T. & Ang, P.C. (1994). Dementia in elderly patients: can the 3R mental stimulation programme improve mental status? *Age and Aging*, 23, 195-199.

McCormack, D. & Whitehead, A. (1981). The effect of providing recreational activities on the engagement level of long-stay geriatric patients. *Age and Aging*, 10, 287-291.