



Aktive Kontrollgruppen als Quelle von Verzerrungen in der Psychotherapieforschung: Eine Metaanalyse anhand von PTSD-Interventionsstudien

Sarah Barkowski¹, Susan Tefikow¹, Jenny Rosendahl¹, Thomas Munder², Jürgen Barth² & Heike Gerger²

¹Universitätsklinikum Jena, Institut für psychosoziale Medizin und Psychotherapie; ²Universität Bern, Institut für Sozial- und Präventivmedizin

Randomisiert-kontrollierte Studien (*randomized-controlled trials*, RCTs) in der Psychotherapieforschung ziehen als Vergleichsgruppen häufig Aufmerksamkeitskontrollgruppen (Placebo) oder bestehende Behandlungsangebote (*treatment-as-usual*, TAU) heran. Für deren Implementierung bestehen jedoch keine allgemeingültigen Richtlinien, die Unterschiede sind dementsprechend groß und ein Risiko für Verzerrungen (*risk of bias*) durch die Kontrollgruppe (KG) besteht. Dies ist sowohl für die Interpretierbarkeit der einzelnen Studien als auch für eine Zusammenfassung in Metaanalysen problematisch. Die vorliegende Studie betrachtet deshalb den Einfluss verschiedener Merkmale des Kontrollgruppendesigns auf die Effektstärkeschätzung.

Risk of bias bezeichnet das Risiko einer systematischen Über- oder Unterschätzung des Behandlungseffekts (Higgins et al. 2011). Bekannte Quellen der Verzerrung stellen allgemeine Designaspekte wie eine fehlende Randomisierung oder eine ungenügende Verblindung des Studienpersonals dar. Studien weisen aber auch auf einen Zusammenhang zwischen dem Kontrollgruppendesign und der Höhe des gemessenen Behandlungseffekts hin und unterstreichen allgemein die komplexe Rolle der Kontrollgruppe in der Psychotherapieforschung. Während ein Effekt der Art der Kontrollgruppe auf die ermittelte Effektstärke bereits seit längerem bekannt ist (Lipsey & Wilson, 1993), hat sich das

Interesse in den letzten Jahren der Rolle einzelner Aspekte des Kontrollgruppendesigns zugewendet. So ist z.B. eine im Vergleich zur Interventionsgruppe abweichende Länge der Behandlung als eine Quelle von Verzerrung vorstellbar. Wampold et al. (2011) und Spielmanns et al. (2010) fanden geringere Effektstärken bei TAU-kontrollierten Studien mit qualitativ hochwertigeren Kontrollgruppe (gemessen unter anderem an der konzeptionellen Nähe zur Psychotherapieintervention und einer vergleichbaren Länge der Behandlung in beiden Gruppen). Baskin et al. (2003) untersuchten den Einfluss der strukturellen Vergleichbarkeit von Placebo-Kontrollgruppen und den Interventionsgruppen auf die Höhe der Effektstärke und berichten, dass bei bestehender Vergleichbarkeit geringere Effektstärken resultieren.

Ausgehend von diesen Ergebnissen überprüft die vorliegende Studie die Hypothese, dass ein hohes Verzerrungspotenzial durch die Kontrollgruppe zu einer Überschätzung der Effektstärke führt. Dafür wurden Aspekte des Kontrollgruppendesigns identifiziert, die theoretisch oder empirisch begründet einen Einfluss auf die Effektstärkeschätzung haben können. Anhand eines Datensatzes randomisiert-kontrollierter Studien zur Wirksamkeit von Psychotherapie bei posttraumatischer Belastungsstörung (PTSD) wurde der angenommene Zusammenhang untersucht.

Methoden

Studien

N = 37 RCTs mit k = 38 Vergleichen zwischen einer Psychotherapiegruppe (CBT: 68%; EMDR: 21%) für erwachsene Patienten mit PTSD und einer Aufmerksamkeits- (71%) oder TAU-Kontrollgruppe (30%).

Publikationsjahr: 1991 - 2009.

Verzerrungsindikatoren

- Vergleichbarkeit der Dosis → Unterschiedliche Behandlungsdauer
- Vergleichbarkeit des Settings → Unterschiede des Settings
- Standardisierung → Unstandardisierte Behandlung
- Nähe zur Psychotherapie → Kein psychologisches Rational
- Glaubwürdigkeit (nur Placebo) → Geringe Glaubwürdigkeit

Hohes Verzerrungspotential - Indikatoren nicht erfüllt

Analysen

Metaanalyse zur Berechnung des mittleren Effekts über alle RCTs (Hedges' g), Metaregressions zur Ermittlung des moderierenden Effekts des KG-Designs.

Ergebnisse

Die mittlere Effektstärke über alle Studien betrug $g = 0.57$ (95% CI [0.37 - 0.76]), $p < .001$. Es zeigte sich eine erhebliche Heterogenität mit $I^2 = 74.4\%$ ($\tau^2 = .25$) und einem signifikanten p -Wert der Q -Statistik ($Q = 144.76$; $d.f. = 37$; $p < .001$).

Erwartungsgemäß wurde eine hohe Variabilität der Studien bezüglich des Kontrollgruppendesigns gefunden. Einige Studien erfüllten alle Indikatoren für geringe Verzerrung, andere keinen. Eine Metaregressionsanalyse mit der Anzahl erfüllter Indikatoren als Prädiktor der Effektstärke ergab einen negativen Zusammenhang (Abb. 1, Glaubwürdigkeit wurde nicht berücksichtigt, da nicht für alle Studien erfasst). Die Ergebnisse nachfolgender Metaregressionsanalysen zur Betrachtung der Indikatoren jeweils einzeln sind in Abb. 2 dargestellt. Signifikante, hypothesenkonforme Effektstärkenunterschiede ergaben sich hinsichtlich der Vergleichbarkeit der Dosis, der Standardisierung und der Glaubwürdigkeit (nur Placebo). Die Heterogenität blieb in allen Regressionsmodellen hoch ($\tau^2 > .20$), aber zeigte eine deutliche Reduktion für Subgruppen von Studien mit einem geringen Verzerrungspotenzial ($\tau^2 < .20$).

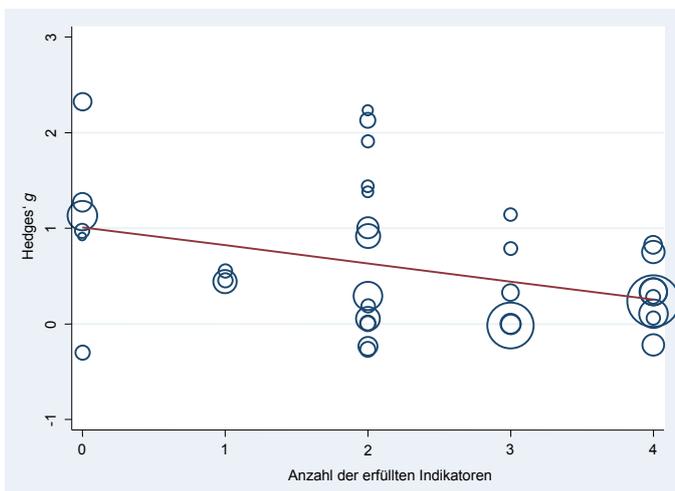


Abb. 1: Die Regressionslinie stellt den negativen Zusammenhang zwischen der Anzahl der erfüllten Indikatoren geringer Verzerrung und der Effektstärke dar (Beta = $-.190$; SE = $.078$; $p = .01$; $R^2_{adj} = 22.41$; $\tau^2 = .248$).

Ergebnisse innerhalb der Subgruppen					Regressionsanalysen zu Subgruppendifferenzen		
Verzerrungsindikatoren	Subgruppen/Verzerrungspotential	Hedges' g (95% CI)	N	τ^2	p-Wert (einseitig)	T^2 (des Modells)	% R^2_{adj}
Vergleichbarkeit der Dosis	hoch	0.91 (0.56, 1.27)	16	.37	.005	.23	28.66
	gering	0.31 (0.13, 0.49)	22	.09			
Vergleichbarkeit des Settings	hoch	0.70 (0.42, 0.98)	27	.38	.06	.30	7.03
	gering	0.27 (0.08, 0.46)	11	.04			
Standardisierung	hoch	1.09 (0.46, 1.72)	6	.46	.02	.26	19.8
	gering	0.46 (0.27, 0.64)	32	.16			
Nähe zur Psychotherapie	hoch	0.68 (0.33, 1.03)	18	.43	.22	.33	-1.85
	gering	0.47 (0.25, 0.69)	20	.15			
Glaubwürdigkeit	hoch	0.73 (0.40, 1.05)	17	.35	.046	.30	8.78
	gering	0.26 (-0.01, 0.52)	10	.08			

Abb. 2: Metaregressionsanalysen zu Effektstärkedifferenzen zwischen Subgruppen von Studien mit Kontrollgruppen deren Verzerrungspotenzial hinsichtlich der einzelnen Indikatoren hoch bzw. gering ist.

Diskussion

Die hohe Heterogenität innerhalb des Studienpools von PTSD-Interventionsstudien deutet auf ein Vorhandensein von Moderatoren des Behandlungseffekts hin. Inwiefern diese Heterogenität durch Aspekte des Kontrollgruppendesigns aufgeklärt werden kann, wurde in der vorliegenden Studie untersucht.

Es ergab sich eine Überschätzung der Effektstärke durch Kontrollgruppen mit hohem Verzerrungspotenzial. Dabei zeigten drei der ausgewählten Indikatoren auch einzeln einen moderierenden Einfluss auf die Effektstärkeschätzung (Vergleichbarkeit der Dosis, Standardisierung, Glaubwürdigkeit (nur Placebo)). Obwohl in keinem der Regressionsmodelle eine deutliche Reduktion der Heterogenität im Gesamtmodell erreicht werden konnte, verdeutlicht die sich ergebende

Homogenität in Subgruppen von Studien mit geringem Verzerrungspotenzial, dass das Kontrollgruppendesign bei der Vergleichbarkeit von Ergebnissen verschiedener Studien eine große Rolle spielt und in Zukunft bei der Zusammenfassung von Studien in Metaanalysen mehr Beachtung erfahren sollte.

Die Ergebnisse reißen sich ein in Befunde neuerer Studien, die annehmen, dass frühere Metaanalysen die Effektstärke in Psychotherapiestudien überschätzen und einen Zusammenhang zur Studienqualität zeigten (Cuijpers et al. 2010). Jedoch ist die Generalisierbarkeit eingeschränkt, da lediglich PTSD-Studien einbezogen wurden. Weitere Studien, die sowohl zusätzliche Populationen als auch die Interaktion mit weiteren Moderatoren, insbesondere der Studienqualität insgesamt (vgl. Higgins et al. 2011), untersuchen, sind daher wünschenswert.

Literatur

- Baskin, T. W., Callen Tierney, S., Minami, T., & Wampold, B. E. (2003). Establishing Specificity in Psychotherapy: A Meta-Analysis of Structural Equivalence of Placebo Controls. *Journal of Consulting and Clinical Psychology, 71*(6), 973-979.
- Cuijpers, P., van Straten, A., Bohlmeijer, E., Hollon, S. D., & Andersson, G. (2010). The effects of psychotherapy for adult depression are overestimated: a meta-analysis of study quality and effect size. *Psychological Medicine, 40*(2), 211-223.
- Higgins, J. P. T., Altman, D. G., & Sterne, J. A. C. (2011). Chapter 8: Assessing risk of bias in included studies. In J. P. T. G. Higgins, S. (Ed.), *Cochrane Handbook for Systematic Reviews of Interventions*. Version 5.1.0: The Cochrane Collaboration.
- Lipsey, M. W., & Wilson, D. B. (1993). The efficacy of psychological, educational, and behavioral treatment: Confirmation from meta-analysis. *American Psychologist, 48*(12), 1181-1209.
- Spielmanns, G. I., Gatlin, E. T., & McFall, J. P. (2010). The efficacy of evidence-based psychotherapies versus usual care for youths: Controlling confounds in a meta-analysis. *Psychotherapy Research, 20*(2), 234-246.
- Wampold, B. E., Burdge, S. L., Laska, K. M., Del Re, A. C., Baardseth, T. P., Fluckiger, C., et al. (2011). Evidence-based treatments for depression and anxiety versus treatment-as-usual: A meta-analysis of direct comparisons. *Clinical Psychology Review, 31*(8), 1304-1312.