

Randomized controlled trials in neurosurgery—how good are we?

K. Schöller · S. Licht · J.-C. Tonn · E. Uhl

Received: 8 February 2009 / Accepted: 9 March 2009 / Published online: 1 April 2009
© Springer-Verlag 2009

Summary

Background The strongest evidence in medical clinical literature is represented by randomized controlled trials (RCTs). This study was designed to evaluate neurosurgically relevant RCTs published recently by neurosurgeons.

Method A literature search in MEDLINE and EMBASE included all clinical studies published up to 30 June 2006. RCTs with neurosurgical relevance published by at least one author with affiliation to a neurosurgical department were selected. The number and characteristics of individual trials were recorded, and the quality of the trials with regard to study design, quality of reporting, and relevance for clinical practice was assessed by two different investigators using a modification of the Scottish Intercollegiate Guidelines Network methodology checklist. Changes of RCT quality over time as well as factors influencing the quality were analyzed.

Findings From the initial search results (MEDLINE $n=3,860$, EMBASE $n=3,113$ articles), 159 RCTs published by neurosurgeons were extracted for final evaluation. Of the RCTs, 62% have been published since 1995; 52% came from the USA, UK, and Germany. The median RCT sample size was 78 patients and the median follow-up 35.7 weeks.

Fifty-two percent of all RCTs were of good, 37% of moderate, and 11% of bad quality, with an improvement over time. RCTs with financial funding and RCTs with a sample size of >78 patients were of significantly better quality. There were no major differences in the rating of the studies between the two investigators.

Conclusions Only a fraction of neurosurgically relevant literature consists of RCTs, but the quality is satisfying and has significantly improved over the last years. An adequate sample size and sufficient financial support seem to be of substantial importance with regard to the quality of the study. Our data also show that by using a standardized checklist, the quality of trials can be reliably assessed by observers of different experience and educational levels.

Keywords Randomized controlled trial · Clinical trial · Evidence-based medicine · Neurosurgery

Introduction

Neurosurgery, on one hand, is a traditional craft: Specialists but also trainees have a rather scholastic view [7, 23], and research as well as development is not always highly regarded [11]. On the other hand, the neurosurgical field faces an increasing number of surgical techniques, technical innovations, medical treatment options, and growing shortage of financial resources [19]. However, despite dealing with traditions and personal experiences, for the good of our patients, it is inevitable to transfer results of evidence-based research into clinical practice to improve our understanding of diseases and their surgical treatment [20].

Randomized controlled trials (RCTs) obviously reflect the gold standard of modern evidence-based clinical

S. Licht · J.-C. Tonn · E. Uhl
Department of Neurosurgery,
University of Munich Medical Center–Grosshadern,
Ludwig-Maximilians-University,
Munich, Germany

K. Schöller (✉)
Department of Neurosurgery,
University of Munich Medical Center,
Grosshadern Marchioninstr. 15,
81377 Munich, Germany
e-mail: Karsten.Schoeller@med.uni-muenchen.de

research [2]. However, only a fraction of neurosurgically relevant topics have been investigated by RCTs [3]. And, furthermore, the quality of relevant RCTs in neurosurgery has been regarded as “in need of improvement” [3–5, 22].

A recent review has presented data on RCTs dealing with neurosurgical procedures [22], but excluded studies on peripheral nerves, physiotherapy, radiotherapy, or drug studies. In contrast, the current study has been designed to assess the quantity, quality, and individual features of the entire body of RCT literature relevant for and published by at least one neurosurgeon as first or senior author, indicating the primary responsibility of a neurosurgeon for the study. Furthermore, non-surgical and medical trials were also evaluated. The quality of the trials was assessed by two different investigators, quality changes over time were documented, and factors having an impact on the quality were extracted. Finally, we sought to analyze whether a standard checklist is reliable when trials are rated by neurosurgeons of different experience and educational level.

Materials and methods

Internet search and selection of articles

A literature search was performed in the medical databases MEDLINE (www.pubmed.com) and EMBASE (www.embase.com) on 30 June 2006. The keywords “neurosurgery” and “clinical trial” were used with no search limits. In a first step, all publications published until that date focusing on issues with relevance for neurosurgical clinical practice were kept for further evaluation; articles with clear focus on anesthesiological or neurological problems were excluded. There was no limitation to the modality investigated by the trial, which means that also non-surgical studies that compared drugs, radiotherapeutic treatments, and others (e.g., physiotherapy) were included. Relevant publications were then reviewed for the first and senior authors' clinical specialization, and only articles with the first or the senior author being affiliated with a neurosurgical department were kept for further evaluation. Letters, editorials, congress proceedings, or technical notes were excluded as well as articles not written in English. The remaining articles were evaluated for whether the criteria for a RCT were fulfilled; otherwise, they were excluded. As a next step both databases were merged, identical manuscripts were identified, and the duplicates were excluded. Multiple publications of the same trial (same groups of patients and same follow-up focusing on different aspects, or double publication of the same data in two journals) were considered only once.

Evaluation of RCT characteristics

All RCTs were examined thoroughly and evaluated by two investigators, a senior neurosurgeon (E.U. = investigator I, 48 years old, 18 years' professional experience) and a neurosurgical resident (K.S. = investigator II, 34 years old, 6 years' professional experience). Initially, the following characteristics of each article were recorded: (1) the database the article was found in (PubMed vs. EMBASE vs. both), (2) year of publication, (3) home country of publishing author, (4) journal, (5) total number of patients in the trial, (6) follow-up period, (7) source of funding (industry vs. independent, e.g., national research organization vs. both) and number of funding sources for a trial, (8) main topic (surgical vs. drug vs. other, e.g., physiotherapy, radiotherapy), (9) topic in detail [vascular, tumor, traumatic brain injury (TBI), hydrocephalus, functional neurosurgery/epilepsy, infections], and (10) single center vs. multicenter trial (MCT), and comparison of results from single centers in a MCT.

Evaluation of RCT quality

In a second step, the quality of each RCT was separately evaluated by both investigators using a modification of the methodology checklist of the Scottish Intercollegiate Guidelines Network (SIGN) [15]. The modified checklist consisted of 12 criteria with each criterion receiving a maximum point score between 1 and 6. The maximum point score per criterion reflected the weight that was determined by the investigators [e.g., the investigators considered that methods applied to minimize the bias (maximum point score=3) were more important than an appropriate and clearly focused question asked in the manuscript (maximum point score=1)]. The sum of the scores per criterion resulted in the RCT point score (maximum 26 points/RCT). A study was classified as bad (B) if 0 to one third of the maximum point score (0–8 points) was reached, as moderate (M) if one third to two thirds of the maximum point score (9–17 points) was reached, and as good (G) if more than two thirds of the maximum point score (18–26 points) was reached. The scores assigned by the two investigators were compared for inter-rater differences. The modified methodology checklist, including the scoring system, is demonstrated in detail in Table 1.

In addition to the assessment of the quality, the change of the quality over time was investigated by generating four groups of time intervals (RCT published \leq 1990, from 1991 until 1995, from 1996 until 2000, and after 2000) and comparing these for overall quality and quality of individual criteria (mean score investigator I and II, respectively). Furthermore, the influence of funding, the follow-up

Table 1 Modification of the SIGN methodology checklist

No.	Criterion		Score
1	Appropriate question addressed by study	Yes	1
		No	0
2	Randomized assignment of subjects to treatment groups	Not reported	0
		Adequate concealment	1
		Concealment by external institute	2
3	Subjects are kept 'blind' about treatment allocation	No	0
		Yes	1
4	Doctors are kept 'blind'	No	0
		Yes	1
5	Investigators are kept 'blind'	No	0
		Yes	3
6	Equality of groups	Not reported	0
		No	3
		Yes	6
7	Relevant outcomes are measured in standard, valid and reliable way	Bad methods	0
		Moderate methods	1
		Good methods	2
8	Intention to treat analysis (ITT)	Not reported/cross over	0
		Yes	1
9	Bias	Yes	0
		No	3
10	Key question is answered	No	0
		Yes	3
11	Results are directly applicable to patient group targeted by this study	No	0
		Yes	2
12	Quality of statistical tests	Bad	0
		Good	1

period, and the sample size on RCT quality was investigated by generating two groups (with/without funding, follow-up, and sample size above and below mean, respectively) and comparing these groups for statistically significant differences of overall RCT quality (investigators' mean score).

Statistical analysis

Statistical analysis was performed using SigmaStat 3.11 Statistical Software (SPSS Science Inc., Chicago, IL). Descriptive analysis was performed using means with standard deviation, medians, and percentages. The Mann-Whitney rank sum test (comparison of two groups) or the Kruskal-Wallis ANOVA on ranks test followed by Dunn's method (comparison of more than two groups) was used to compare non-parametrical variables. Categorical variables were analyzed with the chi-square test. Differences were considered significant at the $p < 0.001$ or $p < 0.05$ levels.

Results

Internet search results and selected articles

The Internet retrieval resulted in 3,860 MEDLINE articles and 3,113 articles in EMBASE; 2,457 articles were excluded from the MEDLINE search results and 1,183 from the EMBASE search results because they were classified as being without neurosurgical relevance. Of the remaining 1,403 MEDLINE articles (=36%) and 1,930 EMBASE articles (=62%), 574 MEDLINE articles and 889 EMBASE articles were excluded because the first or senior author was not a neurosurgeon, leaving 829 MEDLINE articles (=21%) and 1,041 EMBASE articles (=33%). Furthermore, 713 MEDLINE and 925 EMBASE articles were excluded because the article was a letter, an editorial, a congress proceeding, a technical note, or not written in English. One hundred four articles in 12 other languages were identified, and most of these articles were written in Japanese (22 in MEDLINE). After this step, 116 RCTs from MEDLINE (=3%) and also 116 RCTs found in EMBASE (=4%) remained. These articles were merged ($n=23$ identical articles), leaving $n=209$ RCTs. Finally, after exclusion of $n > 1$ articles published with data based on the same trial, 159 RCTs were eligible for further evaluation. Figure 1 gives an overview of the selected articles based on the form of a "decision tree."

Characteristics of RCTs

More than two-thirds of the 159 eligible RCTs have been published since 1995, only 4 studies before 1985, and the first eligible study found was published in 1980 (Table 2). The years 1999 and 2002 were the most productive, with 17 (=11%) of the RCTs having been published, respectively. Regarding the home country of the publishing author, most of the trials (39%) came from the USA, followed by Germany (7%), and the UK (6%).

In absolute numbers, 61 articles were published by authors with affiliation to US departments, 58 by European, and 21 by Asian authors. Eighteen articles were published from authors of other continents (Africa, South America, and Australia). In one article the origin of the author was not given, and the article was therefore excluded from the characteristics analysis.

The number of patients per trial was highly variable (SD:±399.4 patients), ranging from 4 to 4,030 patients; 58% of the trials included fewer than a hundred patients (median: 78 patients). The median sample size did not significantly change over time (99±230 patients≤1995 vs. 70±460 patients>1996). The median follow-up period of all eligible RCTs was 35.7 weeks, with a range from <1 week to 234 weeks (SD:±47.8 weeks). The median follow-up period did not significantly change over time (14±39 weeks≤1995 vs. 24±51 weeks>1996).

Eighty-six trials (55%) did not receive funding or did not mention it in the article; 46% of the funding came from independent sources (e.g., university grants, independent foundations), 39% from the industry, and 15% from both independent and industrial sources. In the majority of these studies (=63%), funding came only from one source. Trials investigating drugs neither received funding more frequently ($p=0.211$) nor did they have a higher number of funding sources ($p=0.360$) as compared to trials investigating surgical techniques or other neurosurgical problems. Furthermore, the type ($p=0.212$) or number ($p=0.450$) of funding was not different in the USA and Europe. There was a statistical trend towards a different number of funding sources ($p=0.084$) in the UK as compared to Germany.

Only one German trial received any source of funding as compared to half of the UK trials.

Only 29% of the eligible RCT were multicenter trials, and the number of MCTs was not found to increase over time (trials≤1995 vs. trials>1995: $p=0.210$). There was a median of 11 centers participating in these studies (SD±22.6 centers), the largest being the Selfotel (CGS 19755) trial for the treatment of severe head injury with 99 centers participating worldwide [10]. Only six MCTs (13% of MCT) reported the results from participating single centers. Table 2 gives an overview of the characteristics of the 159 eligible RCT.

Journals and topics

The 159 RCTs were published in 53 different journals. As non-English articles were excluded, most of the articles were published in the *Journal of Neurosurgery* (38 articles, 24%, impact factor 2007: 1.990), followed by *Neurosurgery* (23 articles, 15%, impact factor 2007: 3.007) and *Acta Neurochirurgica* (11 articles, 7%, impact factor 2007: 1.391). Of the workgroups that published their articles in the *Journal of Neurosurgery*, 54%, 30% and 5% came from the USA, Europe, and Asia, respectively; 48% of the workgroups that published their articles in *Neurosurgery* came from the USA, 35% from Europe, and none from Asia. Most (63%) of the RCTs published in *Acta Neurochirurgica* came from Europe; only one trial from the USA and none from Asia were published in this journal. The two articles with the best impact factor (2007: 52.589) were both published in the *New England Journal of*

Fig. 1 Algorithm of article selection following an Internet search in EMBASE and MEDLINE. Articles were excluded stepwise according to the displayed criteria, and, finally, remaining RCTs of both databases were merged. Only one article per RCT was accepted

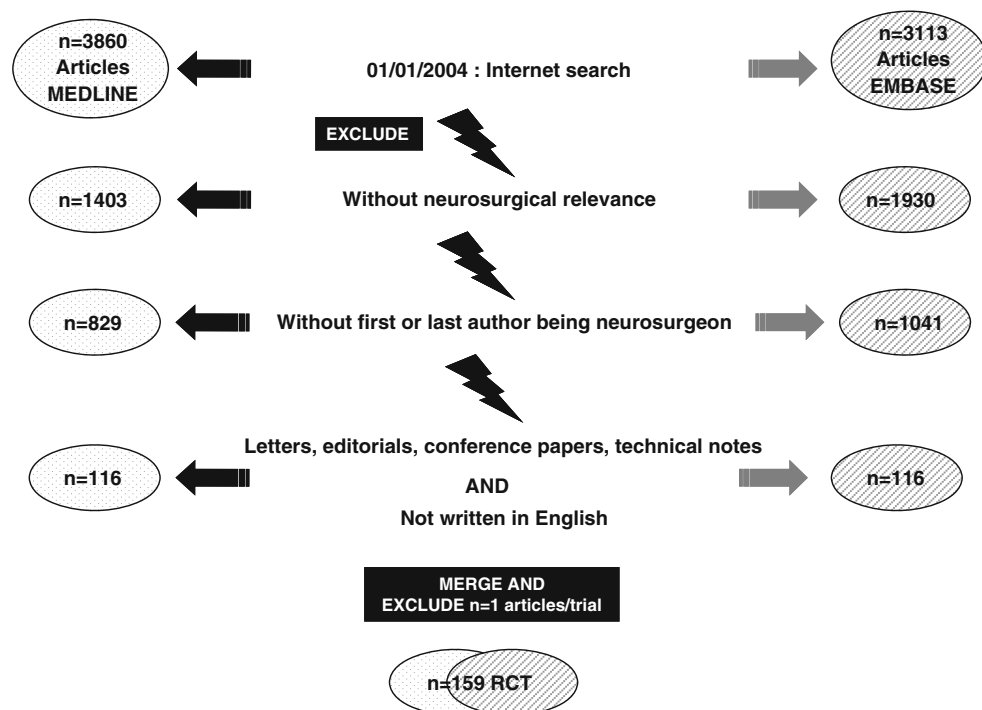


Table 2 Characteristics of RCTs

		n (%)
Year of publication	1980–1985	7 (4)
	1986–1990	21 (13)
	1991–1995	22 (14)
	1996–2000	44 (28)
	2001–2006	65 (41)
Country of publishing author	USA	61 (39)
	Germany	11 (7)
	UK	10 (6)
	Canada	9 (6)
	Japan	9 (6)
	The Netherlands	10 (6)
	China	8 (5)
	Finland	6 (4)
	Sweden	5 (3)
	Others	29 (18)
	Europe	58 (37)
Number of patients/trial	0–100	92 (58)
	100–500	56 (35)
	>500	11 (7)
	Median	78
	Mean	190.6
	Max	4,030
	Min	4
	SD	399.4
	Follow-up period (weeks)	Median
Mean		36
Max		234
Min		<1
SD		47.8
Funding	Funding	71 (45)
	No funding	86 (55)
	Industry	28 (18)
	Independent	33 (21)
	Mixed	11 (7)
	1 source	45 (29)
	2 sources	19 (12)
	3 sources	5 (3)
	>3 sources	4 (3)
Max	9 sources	
Multicenter trials	MCT	46 (29)
	≤1995	10 (6)
	>1996	36 (23)
	No MCT	113 (71)
	Median	11
	Mean	18.7
	Max	99

Table 2 (continued)

		n (%)
	Min	2
	SD	22.6
Comparison of results from single centers	Yes	6 (4)
	Of MCT	(13)
	No	153 (96)

Medicine [12, 13]. Table 3 gives an overview of the ten journals most frequently publishing neurosurgical RCTs.

Most of the trials focused on vascular topics (20%), infection (13%), and traumatic brain injury (13%); 26% of the trials were classified as “other” in the sense of general neurosurgical problems (e.g., local anesthetics prior to placement of skull pins, peripheral nerves, etc). Fifty-four percent of trials investigated drugs, whereas 34% investigated surgical problems; 12% could not be clearly classified as drug or surgical trial and investigated, e.g., radiotherapy or physiotherapy (Fig. 2).

Quality of RCTs

The maximum score to be reached was 26 points. The 159 evaluated RCTs received a mean score of 15.5 ± 6.1 from investigator I and a score of 15.0 ± 6.2 from investigator II. The investigators' mean score was 15.2 ± 6.1 . Investigator I rated 24 (=15%) RCTs to be of bad, 73 (=46%) to be of moderate, and 62 (=39%) to be of good quality. Investigator II rated 27 (=17%) as bad, 81 (=51%) as moderate, and 51 (=32%) as good, respectively. There was no significant difference in the RCT rating between the two investigators ($p=0.435$).

The investigators' rating for the individual RCT criteria differed only for the criterion “methods” ($p=0.001$): Investigator I regarded the methodology in 110 trials as good, in 44 trials as moderate, and in 5 trials as bad, whereas investigator II regarded the methodology as good in 137 trials, as moderate in 20 trials, and as bad in 2 trials, respectively. How the individual quality criteria of the modified SIGN methodology checklist were fulfilled by the 159 evaluated RCTs (investigators' mean score) is displayed in Table 4.

Quality (investigators' mean score) significantly differed ($p<0.05$) between RCTs that were published ≤ 1990 (mean score: 12.2 ± 4.5) compared to RCTs published from 1991– ≤ 1995 (mean score 15.9 ± 4.0), from 1996– ≤ 2000 (mean score: 17.0 ± 4.8), and after the year 2000 (score: 15.0 ± 4.2). There was no significant difference between the time intervals after 1990 (Fig. 3).

Investigation of the individual criteria revealed a significantly different score between the investigated

Table 3 Journals

Journal	Articles (n)
J Neurosurg	38
Neurosurgery	23
Acta Neurochir	11
Surg Neurol	7
Stroke	4
J Neurotrauma	4
JNNP	4
Br J Neurosurg	3
Childs Nerv System	3
Neurosurg Rev	3
Others	59

time intervals for the criteria “Randomization,” “Doctors blind,” “Equality of groups,” “Methods,” “ITT,” “Results directly applicable,” and “Statistics.” In all but the criterion “Doctors blind” the group <1990 received the best score, the group >2000 the worst score, the three more recent groups (>1990) received a significantly better score than the group with articles published before 1990.

RCTs with financial funding (investigators' mean score: 16.8 ± 4.3) had a significantly better quality ($p < 0.001$) than RCTs without funding (investigators' mean score: 13.8 ± 4.5). RCTs with a sample size larger than the median sample size (=78) of all evaluated RCT ($n=80$, score: 17.5 ± 4.1) had a significantly better quality ($p < 0.001$) than RCTs with a sample size ≤ 78 patients ($n=79$, score: 12.7 ± 4.0). RCTs with a follow-up period longer than the median follow-up period of 35.7 weeks ($n=51$)

had the same quality as RCTs with a shorter follow-up period ($n=108$, $p=0.463$).

Discussion

To our knowledge, this is the first study that presents the number, the characteristics, and the quality of the neurosurgical RCTs published in the English language up to July 2006 by neurosurgical-affiliated authors focusing on both surgical procedures and pharmacological interventions that could be retrieved from the Internet databases MEDLINE and EMBASE.

The Internet search for randomized controlled trials revealed different results depending on the medical database that was utilized. EMBASE delivered almost twice as many clinical articles that were relevant for neurosurgical practice and more articles that additionally were written by neurosurgeons. However, the number of neurosurgical RCTs written by authors affiliated with a neurosurgical department was exactly the same, because EMBASE listed more letters, editorials, congress proceedings, and technical notes as well as articles not written in English. Interestingly, 73 of the trials found by MEDLINE and EMBASE were different and only 43 identical. Given this fact, one can derive the recommendation to search both of these databases if attempting a thorough literature search, at least for neurosurgical issues. Furthermore, one should be aware that EMBASE presents more clinical publications that are relevant for neurosurgeons.

RCTs accounted for 4% of the clinical neurosurgical literature found in EMBASE and for 3% of the neurosur-

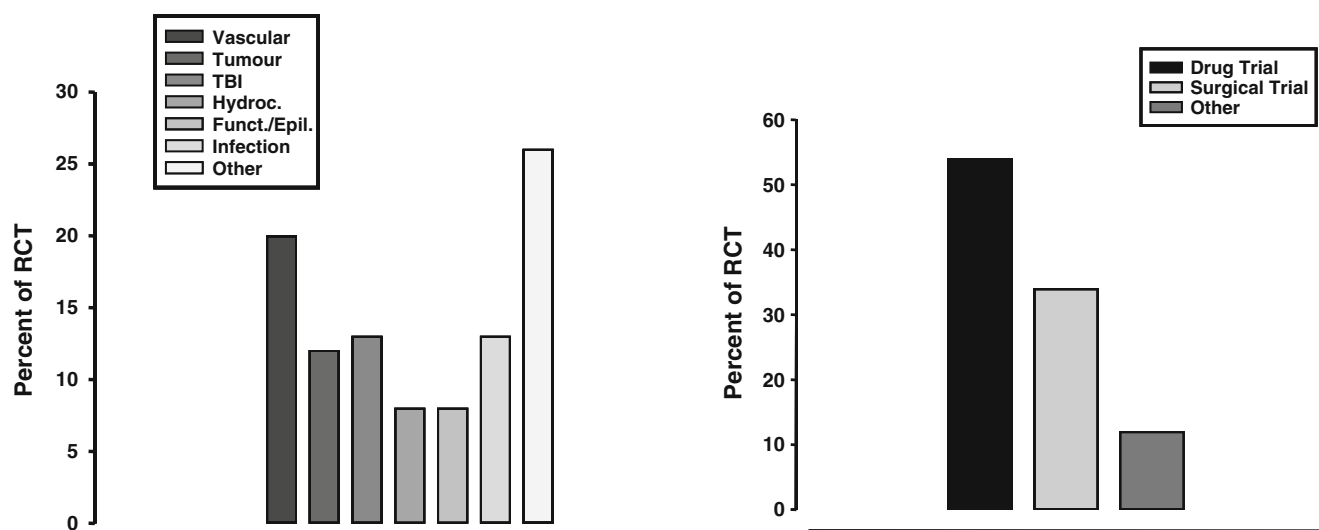


Fig. 2 Topics covered by the 159 eligible RCTs. The most frequent publications were on vascular neurosurgery, infection, and traumatic brain injury. RCTs focusing on spinal neurosurgery comprised only a

small fraction and were included in “other.” Most of the trials compared drugs rather than neurosurgical procedures

Table 4 Results of RCT quality rating—individual criteria (all trials, n=159)

No.	Criterion		n (%)
1	Appropriate question	No	29 (18)
		Yes	130 (82)
2	Randomization	Not reported	73 (46)
		Adequate concealment	68 (43)
		Concealment by external institute	18 (11)
3	Subjects blind	No	90 (57)
		Yes	69 (43)
4	Doctors blind	No	107 (67)
		Yes	52 (33)
5	Investigators blind	No	97 (61)
		Yes	62 (39)
6	Equality of groups	Not reported	35 (22)
		No	25 (16)
		Yes	99 (62)
7	Methods	Bad	3 (2)
		Moderate	32 (20)
		Good	124 (78)
8	ITT	Not reported/cross over	104 (65)
		Yes	55 (35)
9	Bias	Yes	107 (67)
		No	52 (33)
10	Key question answered	No	25 (16)
		Yes	134 (84)
11	Results directly applicable	No	71 (45)
		Yes	88 (55)
12	Quality of statistics	Bad	24 (18)
		Good	141 (82)

gical clinical trials found in MEDLINE. This is a dimension that is consistent with the findings of Uhl and coworkers [21], who identified 2.8% RCTs in the most significant general surgical journals, and with the study of Curry and coworkers who identified 1% RCTs in pediatric surgical literature [1]. Solomon and colleagues [16–18] found that surgical RCTs were performed by surgeons in only one third of trials and assumed a lack of expertise by surgeons in clinical trials, methodological problems peculiar to surgical trials, or a need for adoption of other research designs to assess surgical therapies [16]. This is underlined by the results of our study where only 50–60% of the neurosurgically relevant clinical trials have been conducted by authors affiliated with a neurosurgical department. Furthermore, both Salomon et al. and Johnson et al. [7, 16] mentioned that it is more difficult to find commercial funding for a surgical trial than for a drug trial, which, in

most of the cases, is funded by a pharmaceutical company. In the current study, however, we were not able to demonstrate a difference in funding between drug trials and surgical trials. In this context it may play a role that, similar to the study of Vranos et al. [22], in more than half of the trials the funding source was not reported, and it was therefore impossible to determine whether there was any funding in these cases. Nevertheless, it is interesting in this context that funded RCTs were of significantly better quality than RCTs without funding. This could be due to the necessity to write a good study protocol if an investigator applies for funding. Furthermore, several assessments of the study protocol by external evaluators might be required before the financial support is assigned.

Most of the 159 neurosurgical RCTs identified with the help of our algorithm have been published since the mid 1990s. The slowly increasing number of RCTs might be a result of the growing awareness for the need of reliable data that, in the last years, have been pushed forward by the evidence-based medicine (EBM) movement [2, 18]. The finding of our study that the most actively publishing workgroups came from the USA, Germany, and the UK might be limited because a significant number of articles from Asia and Europe written in the native language were excluded. The data of these studies that may be of excellent quality are often not recognized by the neurosurgical community, which is focused on a small spectrum of journals written in English. Nevertheless, Uhl and coworkers mentioned [21] that a growing number of articles are published in English due to the “impact factor restraint.”

Assessment of the quality of the RCTs by two investigators with different experience in the field of

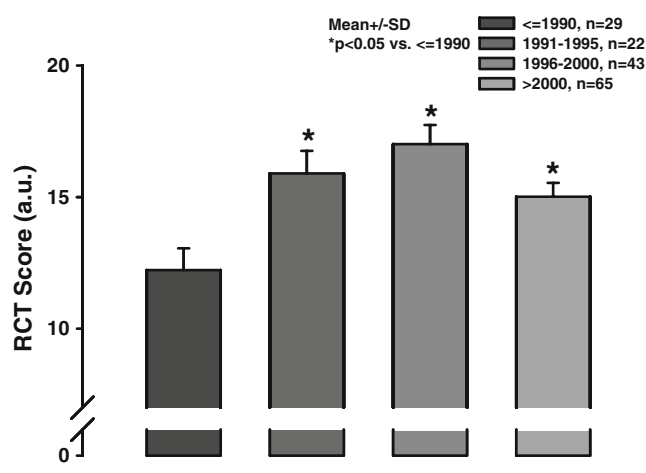


Fig. 3 Quality of RCTs in different time periods (mean score of both investigators). The quality of neurosurgical RCTs improved significantly over time with the best results in the period from 1996 through 2000 (mean \pm SD, * $P < 0.05$ vs. ≤ 1990 , one-way ANOVA and Student-Newman-Keuls method)

neurosurgery aimed at maximum validity. Using a standardized checklist modified after the SIGN criteria for RCTs, the separate evaluation yielded quite similar results for most of the investigated criteria, and the overall score was not significantly different. However, the more experienced investigator (= investigator I), in general, assigned a better score for RCT quality. This can be attributed to the larger experience of investigator I with medical research/medical literature as well as with the implementation of clinical trial results into daily clinical neurosurgical practice. Moreover, differences between the investigators that occurred with assessment of the methodology are not surprising since these criteria are less precise. Overall, the SIGN methodology checklist was found applicable for the assessment of RCT quality and can be recommended by the authors.

For the evaluation of an individual RCT, however, one has to take into account which criterion of RCT quality could have been addressed and which could not have been addressed due to a certain trial design. For example, it is not always possible to carry out a double-blind study if a surgeon has to use a certain technical device for an operation in one group of patients and not in the other group. However, similar to Vranos and colleagues [22] who focused on neurosurgical trials that involved surgical procedures, we found that allocation concealment was infrequently ensured and that even assessors of outcome in most of the cases were not blinded—both methodological limitations that, in most of the cases, could have been overcome. And there are quite a number of surgical trials that demonstrate how patient blinding can be achieved with an intelligent trial design [9, 14, 24].

Despite the criticism of neurosurgical RCTs, it has to be pointed out that almost half of the trials were of good quality and almost 90% of at least moderate quality. Also, the number of multicenter trials has been increasing during the last years. Furthermore, the overall quality is similar to other disciplines [6, 8] and is, at least slightly, improving over time. This change seems to be an effect of a better randomization scheme, a better methodology, a better equality of treatment groups, a better statistical analysis, as well as a better applicability of RCTs to clinical practice, which is probably a key criterion of a good trial design.

References

- Curry JI, Reeves B, Stringer MD (2003) Randomized controlled trials in pediatric surgery: could we do better? *J Pediatr Surg* 38:556–559. doi:10.1053/jpsu.2003.50121
- Guyatt GH, Sackett DL, Sinclair JC, Hayward R, Cook DJ, Cook RJ (1995) Users' guides to the medical literature IX. A method for grading health care recommendations. Evidence-Based Medicine Working Group. *JAMA* 274:1800–1804. doi:10.1001/jama.274.22.1800
- Haines SJ (1979) Randomized clinical trials in the evaluation of surgical innovation. *J Neurosurg* 51:5–11
- Haines SJ (1983) Randomized clinical trials in neurosurgery. *Neurosurgery* 12:259–264. doi:10.1097/00006123-198303000-00001
- Haines SJ (2001) History of randomized clinical trials in neurosurgery. *Neurosurg Clin N Am* 12:211–216
- Hill CL, LaValley MP, Felson DT (2002) Secular changes in the quality of published randomized clinical trials in rheumatology. *Arthritis Rheum* 46:779–784. doi:10.1002/art.512
- Johnson AG, Dixon JM (1997) Removing bias in surgical trials. *BMJ* 314:916–917
- Kyriakidi M, Ioannidis JP (2002) Design and quality considerations for randomized controlled trials in systemic sclerosis. *Arthritis Rheum* 47:73–81. doi:10.1002/art1.10218
- Majeed AW, Troy G, Nicholl JP, Smythe A, Reed MW, Stoddard CJ, Peacock J, Johnson AG (1996) Randomised, prospective, single-blind comparison of laparoscopic versus small-incision cholecystectomy. *Lancet* 347:989–994. doi:10.1016/S0140-6736(96)90143-9
- Morris GF, Bullock R, Marshall SB, Marmarou A, Maas A, Marshall LF (1999) Failure of the competitive N-methyl-D-aspartate antagonist Sefotel (CGS 19755) in the treatment of severe head injury: Results of two phase III clinical trials. *J Neurosurg* 91:737–743
- Ostertag CB (2001) Measuring the importance of scientific results in neurosurgery. *Acta Neurochir Suppl (Wien)* 78:185–188
- Penn RD, Savoy SM, Corcos D, Latash M, Gottlieb G, Parke B, Kroin JS (1989) Intrathecal baclofen for severe spinal spasticity. *N Engl J Med* 320:1517–1521
- Perl TM, Cullen JJ, Wenzel RP, Zimmerman MB, Pfaffler MA, Sheppard D, Twombly J, French PP, Herwaldt LA, Bentler S, Brown K, Bushnell W, Coffman S, Finlay J, Forbes R, Hintemier M, James C, Helgerson B, Preston A, Swift J (2002) Intranasal mupirocin to prevent postoperative *Staphylococcus aureus* infections. *N Engl J Med* 346:1871–1877. doi:10.1056/NEJMoa003069
- Richter HP, Kast E, Tomczak R, Besenfelder W, Gaus W (2001) Results of applying ADCON-L gel after lumbar discectomy: the German ADCON-L study. *J Neurosurg* 95:179–189
- Scottish Intercollegiate Guidelines Network (SIGN) SIGN 50 (2005) A guideline developers' handbook. Methodology Checklist 2: Randomised Controlled Trials www.sign.ac.uk
- Solomon MJ, Laxamana A, Devore L, McLeod RS (1994) Randomized controlled trials in surgery. *Surgery* 115:707–712
- Solomon MJ, McLeod RS (1995) Should we be performing more randomized controlled trials evaluating surgical operations? *Surgery* 118:459–467. doi:10.1016/S0039-6060(05)80359-9
- Solomon MJ, McLeod RS (1998) Surgery and the randomised controlled trial: past, present and future. *Med J Aust* 169:380–383
- Thomas DG, Kitchen ND (1994) Minimally invasive surgery. *Neurosurgery*. *BMJ* 308:126–128
- Troidl H (1998) Conceptual and structural conditions for successful clinical research. *Langenbecks Arch Surg* 383:306–316. doi:10.1007/s004230050138
- Uhl W, Wente MN, Buchler MW (2000) Surgical clinical studies and their practical realization. *Chirurg* 71:615–625. doi:10.1007/s001040051113

22. Vranos G, Tatsioni A, Polyzoidis K, Ioannidis JP (2004) Randomized trials of neurosurgical interventions: a systematic appraisal. *Neurosurgery* 55:18–25. doi:[10.1227/01.NEU.0000126873.00845.A7](https://doi.org/10.1227/01.NEU.0000126873.00845.A7)
23. Winkler JT (1987) The academic celebrity syndrome. *Lancet* 21:450. doi:[10.1016/S0140-6736\(87\)90151-6](https://doi.org/10.1016/S0140-6736(87)90151-6)
24. Wyler AR, Hermann BP, Simes G (1995) Extent of medial temporal resection on outcome from anterior temporal lobectomy: a randomized prospective study. *Neurosurgery* 37:982–990

Comments

Every single neurosurgeon should read and study the randomized trials herself/himself and he/she should evaluate the value individu-

ally. One should be aware that not every trial gives level I evidence, and all are not drawn according to trial guidelines.

V. Benes
Prague, Czech Republic

The actual results reflect the state of RCTs during the EBM hype. The total number of RCTs is very small, actually too small to have a substantial impact on the developments in our field. The current situation is even less promising for RCTs because the premise of equipoise for studying neurosurgical issues does not work anymore. The current distribution of information and the propagation of guidelines leave no database anymore for surgical trials under the equipoise premise. Only emerging new techniques can be compared with traditional ones under these preconditions.

H-J. Steiger
Dusseldorf, Germany