Endoscopic spine surgery outcome studies

Endoscopic spine surgery may be questioned for its medical necessity and effectiveness when compared to other forms of spinal surgeries. The call for clinical evidence to justify the need for capital equipment purchases, disposables, and additional training often follows and is echoed by the repetitive question why endoscopy is better and more cost effective than traditional open or other types of translaminar minimally invasive spine surgery. What is evident though, is the pushback by payers purely against any advancement in spine care because of escalating costs. Lack of clinical evidence is the number one cited reason why coverage for proposed endoscopic spine care is denied by deeming it experimental and medically not necessary. Clinical coverage- and treatment guidelines written by payers and surgical societies are centered around available high-grade evidence published in the peer-reviewed literature. Most outcomes studies in endoscopic spine surgery are level III retrospective case series published by pioneers of the procedure including the senior author of this editorial—Anthony Yeung, MD (1). Level I, and II prospective randomized trials are few far and between. However, a few high-grade studies comparing outcomes with endoscopic versus microsurgical decompression have been published.

Ruetten (2-4), Komp (5,6) and their team have taken the outcome research approach to validate endoscopic surgery techniques seriously. In 2009, Ruetten et al. published the results on surgical treatment for lumbar lateral recess stenosis with the full endoscopic and interlaminar approach versus conventional microsurgical technique (2). This prospective randomized controlled trial on 161 patients showed similar clinical results in the “full endoscopic group” and the microsurgical group when analyzing the German version of the North American Spine Society instrument and the Oswestry low back pain disability questionnaire. In another study, the same team of authors included 178 patients and reported complete relief of leg pain in 82% of the patients at two-year follow-up. An additional, 14% only had occasional pain. The clinical results were similar between traditional microdiscectomy and full endoscopic technique. Even the reported recurrence rates (6.2%) were no differences between the conventional microdiscectomy and full endoscopic discectomy. Ruetten et al. suggested that there were significant advantages with the full-endoscopic techniques including less back pain, improved rehabilitation, fewer complications, and less traumatization, however, offered few objective data to support his clinical observations. Moreover, the authors did not explain whether any patients did not complete the minimum follow-up in their randomly assigned group. The authors concluded that clinical outcomes are equal at two-year follow-up were comparable between the two surgical treatments of a herniated disc (2). In 2011, the authors’ findings were corroborated by a follow-up study on 87 patients with recurrent herniation after conventional discectomy who underwent full-endoscopic or microsurgical intervention (6). At two-year follow-up years and utilizing the same clinical outcome measures, the authors reported a 79% success rate.
with patients no longer having had any leg pain, and a recurrence rate was 5.7% with no difference between the groups.

In 2013, Birkenmaier et al. performed a metaanalysis of comparative controlled clinical trials on endoscopic and microsurgical standard procedure (7). His review focused on full-endoscopic including interlaminar and transforaminal approaches for all spinal regions. His literature analysis of a PubMed and Embase search considering entries up to January 2013 included a total of 504 studies. Four randomized controlled trials (RCTs) and one controlled studies (CS) were finally identified as eligible for evaluation. Birkenmeier analyzed these five manuscripts concerning randomization, inclusion and exclusion criteria, clinical outcomes, and complications. Four studies consistently showed shorter operating times, less blood loss, less operative site pain, and faster postoperative rehabilitation, shorter hospital stay, sooner return to work with the endoscopic techniques when compared to the microsurgical techniques. Clinical outcomes were found to be similar between the endoscopic and the microsurgical methods in any of the trials. All five studies had fewer complications with the endoscopic technique, and this was statistically significant in 2 of the reviews. One study showed a lower rate of revision surgeries requiring arthrodesis with the endoscopic procedure.

In 2019, Kong et al. prospectively randomized patients with lumbar disc herniation and lateral recess stenosis to either endoscopic lumbar discectomy or microsurgical laminotomy technique (8). This small study of 40 patients showed similar clinical outcomes using the Oswestry Disability Index (ODI) and visual analog scales (VAS) for back pain and leg outcome measures for both treatments at two-year follow-up. Another open-label randomized single-center study published in 2018 by Limin Rong’s team at the Third Affiliated Hospital of Sun Yat-sen University of Medical Sciences in Guangzhou China attempted to compare percutaneous transforaminal endoscopic discectomy to mid-line microendoscopic discectomy (9). A computer-generated code did randomization of the 153 participating patients. Outcome measures were ODI, VAS for back and leg pain, and Medical Outcomes Study 36-Item Short-Form Health Survey bodily pain and physical function scales (SF-36), and EuroQol Group’s EQ-5D. One-year follow-up data were available in 89.5%. Additionally, length of surgery and hospital stay, time to mobilization, cost of surgery, and total hospital cost, complications-, and reoperations rates were recorded. The authors reported no difference in primary and secondary outcome measures between the treatment (P=0.05). Subrogating the data obtained in the endoscopic surgery group by location of the disc herniation the authors found less favorable clinical outcomes with smaller ODI reductions with medially located disc herniations at 1 week (P=0.027), 3 months (P=0.013), 6 months (P=0.027), and 1 year (P=0.028) compared with the paramedian subgroup. Far lateral disc herniation treated with translaminar microsurgical decompression were associated with significantly lesser ODI score improvements at 3 months (P=0.008), 6 months (P=0.028), and 1 year (P=0.028). These results suggest that outcomes are best with either technique if the compressive pathology is approached directly. Clinical outcomes may deteriorate if the compressive pathology is more distant from the primary access. The authors reported similarly high complication rates of 13.75% in the endoscopic surgery group and 16.44% in the microsurgical tubular retractor group (P=0.642). The authors concluded that at 1-year follow-up, the superiority of endoscopic surgery approach regarding clinical outcomes could not be demonstrated and that it was not necessarily safer either (9).

Does it really make no difference?

One is left to wonder whether that is really all there is to it: does spinal endoscopic decompression procedures just render similar clinical outcomes as traditional open or other forms of translaminar minimally invasive spinal surgeries? Put differently: clinical endoscopic surgery outcomes are no worse than other types of spine surgeries but are they better for the patient? The traditionally trained spine surgeons may question why to get into endoscopic spine surgery if there is no clear advantage in clinical outcomes? Obviously, something is missing here to describe the benefits of endoscopic spine surgery in a way that matter to patients and their families more than a score on a functional outcome scale or a category assignment rated two years postoperatively—the classic minimum follow-up often reported in cross-sectional comparative case-control studies. The question arises whether these traditional outcome tools including the visual analog (VAS) leg and back pain score (10), the ODI (11), Roland Morris score, the short-form (SF) SF-12 and SF-36 are sensitive enough to detect the factors impacting patient satisfaction and clinical outcomes with endoscopic surgery in an ambulatory surgery center setting (ASC)? These traditional outcome
tools do not adequately reflect the more favorable patient perception with the endoscopic surgery when compared to the types of translaminar surgeries it is trying to replace and also do not measure the benefits of staging treatment options by targeting the predominant pain generator as opposed to treating all sources of pain currently focused on fusion. This disconnect between traditional outcome tools and patient self-reported outcome measures (PROMs) has been well recognized for quite some time and led to the implementation of such PROMs (12). Contemporary clinical outcome studies employing PROMs attempt to better understand the benefit and durability of modern spine care on patients’ functioning, return to work, utilization of other health services, decreased narcotic utilization and improved social reintegration (12). Hence, investigating clinical questions specific to endoscopic spine surgery are more likely to reveal meaningful differences between treatment arms and aid in the statistical analysis of confounding factors affecting patient satisfaction and utilization.

**Why are there so few high-level evidence spine outcome studies?**

The short answer to this question is straightforward: High-level prospective randomized controlled clinical outcome studies in spine surgery are expensive and difficult to carry out. The Spine Patient Outcome Research Trial (SPORT) attempted it and ran from 2000 to 2004 by enrolling 2,500 patients. Initially, it reported on clinical outcomes with surgical versus non-surgical treatment for a herniated disc (13). Later, other indications, including spinal stenosis and spondylolisthesis, were analyzed similarly (14). This multicenter clinical study was carried out in 13 participating spine clinics in 11 states and was supported by the United States National Institute of Health (NIH). Ultimately, the SPORT trial provided inconclusive clinical evidence on the superiority of surgical treatment over non-operative treatment for these indications based on the results of the intent-to-treat analysis. The biggest problem the authors of the SPORT studies ran into was the crossover from one treatment arm into another because patients did not like their randomization assignment. Fifty percent of patients disagreed with their surgeons on being randomized to surgery, and thirty percent of patients did not want non-surgical treatment (14). Randomization and double blinding is practically impossible when procedures are performed only local anesthesia. Patients in pain with significant disability randomized to an ineffective non-operative treatment arm instantly know if their treatment is inferior because of the persistence of pain. Unlike in prospective randomized comparative drug trials, an additional test to detect responsiveness to the treatment or absence thereof is not necessary for spine surgery if the patients’ pain is either not subsiding or persisting. Therefore, patients crossing over to other treatment arms or leaving the study altogether if their symptoms are not successfully treated remains the main challenge with randomization protocols in spine surgical outcome research.

Another common problem is the general lack of adequate control groups in clinical spine outcome studies. Assigning patients to a placebo control group without any treatment to mostly let the natural history of the underlying degenerative disorder play itself out is not only hard to do but ethically inappropriate in most study scenarios. Institutional Review Boards (IRB) may not grant such a study design. Millum and Grady reviewed the contemporary ethical analysis and international ethical guidelines published by the Council of International Organizations of Medical Sciences’ (CIOMS) (15). They also discussed the International Ethical Guidelines for Biomedical Research Involving Human Subjects regarding the design of controlled clinical studies (15). CIOMS allows for placebo controls in randomized trials when: (I) there is no effective treatment; or (II) if denying treatment has trifling risks and poses no harm to participants. Placebo may also be deemed appropriate if the trial is intended to find a cure that does not exist, and participation in the trial does not result in relinquishment of regular treatment. In spine surgery, these latter two methodological reasons do often not apply in highly industrialized Western countries and the need for a placebo control to demonstrate the efficacy of a new treatment does not necessarily justify it. While there is widespread agreement that the absence of effective treatment and insignificant risk exposure on the study participants’ part justifies placebo controls, the discussion about the interpretation of the CIOMS recommendations and how they apply under specific local conditions—particularly in third world countries with poorly developed health care systems—remains controversial when methodological reasons are cited to justify the increased risk exposure to study participants. In other words, how much pain would be acceptable for a patient with a symptomatic degenerative spine condition to justify his or her randomization to a placebo control group? It quickly becomes apparent that these discussions are mostly academic and highly impractical for orchestrating
meaningful spine care outcome research. For endoscopy specifically, it comes down to comparing the effectiveness of competing treatments, which means we are back to straightforward level III case-control studies. The latter are conducive to be carried out in busy highly specialized niche private practice settings where large numbers of patients with the condition to be treated are preferably referred and can be quickly enrolled between different sites in the various study arms.

A practical way to study spinal endoscopy outcomes

Cross-sectional retrospective or longitudinal prospective case series studies with average two-year follow-up comparing clinical outcomes with various surgical treatments has been the preferred clinical study in spine care outcome research for all the reasons discussed above. This type of clinical study suffers from a number of limitations ranking it typically as level III (cohort and case-control studies, or systematic review of these studies), level IV (case series), or level V (expert opinion, case report or clinical example; or evidence based on physiology, bench research or “first principles”) (16). Besides the limitations of the commonly used outcome tools, such low-level outcome studies typical of spine care analysis suffer from additional limitations affecting the interpretation of patient selection criteria, and the determination of preoperative prognosticators of favorable clinical outcome with the endoscopic spine surgery. Additional problems may arise due to insufficient follow-up since not all patients enter and exit the study at the same time. Enrolling a patient into a study may also not be identical with the beginning of treatment. It is difficult to adjust for these variables in particular with endoscopic decompression surgery because of its outpatient nature. The surgery is typically short and uneventful. Patients are often seen in follow-up within the immediate postoperative period and then often leave the practice because of the resolution of their symptoms and an uncomplicated postoperative course. Complex surgeon-patient interactions to direct wound care or any other type of aftercare, including physical therapy, or interventional pain management are rarely needed. Hence, patients only tend to return to the office for an unrelated new problem or if symptoms recur. Therefore, significant gaps in follow-up may occur, and the most valuable information obtained from the patient returning to clinic is his or her functional status at the time of reevaluation. If the treatment stopped benefitting the patient (endpoint variable)—the event—it is a measure of the durability of the treatment effect—or putting it in statistical analysis terms—it is the cornerstone of survival analysis in clinical scenarios with incomplete observations as described for spinal endoscopy.

The construction of survival curves was the brain child of Edward L. Kaplan and Paul Meier, who in 1958 developed a statistical analysis method of estimates of survival data in studies with differing survival times (durations of treatment benefit) because endoscopic spine patients often enter the study at different times and have variable duration of clinical benefit from (times-to-event) (17). Frequently, not all spinal endoscopy patients are available for follow-up or remain in the study and are censored (excluded from calculations) because the exact survival time cannot be determined. Another common censoring scenario occurs at the end of the study when patients are still doing well and are being eliminated from the survival calculations because it is unknown how long into the future their treatment benefit would have lasted. The time from entering the study to censoring is called the serial time in contrast to the calendar or secular time. Calendar time describes the traditional design of clinical trials.

Deliberating the specific considerations regarding an outpatient endoscopic spine practice described above, it is the opinion of the editors of this special focused issue on staged management of painful spinal conditions that the construction of Kaplan–Meier (KM) survival curves is better suited for the clinical outcome analysis and derivation of prognosticators of favorable clinical outcomes in endoscopic spine surgery patients than other types of descriptive cross tabulation- or analysis of variance statistics performed at the end of a case-control cohort study. Instead, each subject is characterized by the serial time, the status at the end of their serial time (event occurrence or censored), and their study group assignment. The graphic visualization of the KM survival curves is the result of the construction of survival time probabilities for which purpose the serial times for individual subjects are arranged from the shortest to the longest, without regard to when they entered the study. Thereby, all endoscopic spine patients within a treatment group start the analysis at the same point. They are surviving until (I) the treatment benefit has disappeared—the event of interest, or (II) they have been censored. In other words, only the duration of known survival is measured.
Durability analysis of the endoscopic spinal surgery procedure

The question of whether the endoscopic surgical treatment is more beneficial than other types of minimally invasive or open traditional spine surgery is complicated, but can at least in part can be answered by examining the durability of the treatment benefit defined as continued self-reporting of favorable PROM outcomes by patients without utilization of any other service to manage the underlying degenerative disease process. The KM curve construction is made for this durability (survival) analysis. Understanding what goes into the curve construction is essential to interpreting the KM analysis of endoscopic spine surgery outcome investigation. The survival intervals are plotted in horizontal lines along the x-axis. The length of the lines represents the duration of survival until the end of the endoscopic surgery benefit. The vertical lines connect the horizontal lines to give the appearance of a curve, but the distance of the vertical lines represents the change in the cumulative probability of durability of the endoscopic spine surgery benefit with increasing follow up plotted on the y-axis. In other words, the KM curves are non-continuous. They are step-wise estimates of endoscopic treatment durability. These estimates for the entire study group are most accurate after the minimum follow-up—typically two years for most spine outcome research studies—because the status each patient is known. As soon as the first patient is censored, it becomes an estimate. The accuracy of these estimates deteriorates further on the right side of the curve as fewer patients remain in the study group. The more patients are censored early in the study; the less reliable is the survival curve. Therefore, extrapolation to predict future functioning with endoscopic procedures should be avoided.

Differences between KM curves produced by different endoscopic treatments or analysis of responsiveness to treatment by various endoscopically visualized pain generators can be tested for statistical significance. This is commonly done with the log-rank test. It calculates the chi-square ($\chi^2$) for each event time for each group and sums the results. These are added to obtain the ultimate $\chi^2$ to compare the full curves of each group. For practical reasons, the two- or five-year fifty-percentile survival percentages for the survival curves are often reported so that curves can be readily compared. It goes without saying that the editors have merely described the basics of Kaplan-Meier survival curves with the intent of illustrating its utilization in clinical research of the benefit of endoscopic spine surgery which is often hampered by practical hurdles described above. Despite their limitations of increasingly inaccurate estimates of the duration of treatment benefit as the outcome study continues, KM curves provide an easy-to-understand visualization of the survival effect between various study groups.

An exemplary scenario

One of the articles in this special focus issue compared the clinical outcomes with the transforaminal directly visualized endoscopic surgical versus the non-visualized percutaneous laser interventional disc decompression by performing Kaplan-Meier survival analysis of the duration of the treatment benefit to define the clinical role of these two treatments better (18). This study on a total of 248 patients investigated the impact of several confounding factors on the durability of the treatment benefit of the two procedures. For example, do patients with advanced degenerative changes have less favorable outcomes in the long-run? Reduced posterior disc and lateral recess height are accepted variables often examined in these common clinical scenarios. One would expect that advanced degenerative changes are correlated with shorter durability of the treatment benefit. However, the KM analysis on a series of 248 patients divided into two groups of endoscopic surgical versus percutaneous laser decompression demonstrated the exact opposite at a statistical significance level of $P<0.0001$ in that study (Table 1). The survival curves are shown in Figure 1A and B visually demonstrate how the survival or durability of the treatment benefit from the two entirely different treatments are seemingly affected by the lateral recess height. While further analysis is necessary to understand the significance of other confounding factors to explain this relationship, possible explanations include that the advanced bony and soft tissue decompression achieved by the transforaminal endoscopic surgery under direct visualization can accomplish a better decompression in patients with advanced degenerative changes than the laser (Figure 1A; Table 1). Another possible explanation is that patients with preserved lateral recess height due to less advanced degeneration of the spinal motion segment have a greater potential for progressive disc collapse with the endoscopic procedure as a function of the underlying disease rather than the treatment, and, hence, have shorter durability of the treatment benefit than seen in patients with advanced degeneration, where the vertical travel distance of a collapsing spinal motion segment has nearly
been exhausted. In comparison, patients with less advanced degeneration of the spinal motion segment as evidenced by preserved lateral recess height actually did have a marginal increase in median durability (survival) of the treatment benefit from the laser as indirect decompression due to thermal shrinkage relies on the presence of such disc tissue (Figure 1B; Table 1). This illustrative example shows how KM curve reconstruction and survival analysis can aid in a better understanding of patient selection for spinal endoscopy surgery and how it compares to different treatments.

**Conclusions**

Outcome research in spine surgery is hampered by cross-over problems which plagued well-designed prospective research.
randomized controlled clinical trials. The future of endoscopic spine surgery adoption by payers, professional societies, and governmental review boards will likely hinge in part on providing higher-grade clinical evidence. One of the considerations when advocating to replace traditional open or other types of minimally invasive translaminar with endoscopic spine surgeries is its durability. Spinal endoscopy stands a good chance of gaining more traction in such way mainly if efficacious spine care can be provided at a lower long-term societal cost due to longer duration of treatment benefits, lower utilization of other concomitant treatment services, reduced rate of iatrogenic problems requiring follow-up revision surgeries, and better return to work rates. Proving these empirical benefits well known to most spine surgeons who perform endoscopic surgery at a higher-grade clinical evidence level is critical to improving its acceptance into mainstream and the implementation of reimbursement structures that incentivize surgeons, ASCs and hospitals to break with the old ways. Statistical analysis is a tool to achieve this goal, but it may be difficult to understand for most practicing spine surgeons. In the opinion of the two editors of this special focus issue on staged management of common pain generators in the degenerative spine, plotting of Kaplan Meier survival curves may prove useful to illustrate and communicate the effectiveness and perhaps even the superiority of spinal endoscopy over open surgery in an easy-to-understand visual manner.

**Acknowledgments**

None.

**Footnote**

*Conflicts of Interest:* The authors have no direct (employment, stock ownership, grants, patents), or indirect conflicts of interest (honoraria, consultancies to sponsoring organizations, mutual fund ownership, paid expert testimony). The authors are not currently affiliated with or under any consulting agreement with any vendor that the clinical research data conclusion could directly enrich. This manuscript is not meant for or intended to endorse any products or push any other agenda other than the associated clinical outcomes with lumbar endoscopic surgery. The motive for compiling this clinically relevant information is by no means created and/or correlated to directly enrich anyone due to its publication. This publication was intended to substantiate contemporary endoscopic spinal surgery concepts to facilitate technology advancements. A Yeung designed and trademarked his inside-out YESS™ technique and receives royalties from the sale of his inventions. Indirect conflicts of interest (honoraria, consultancies to sponsoring organizations are donated to IITS.org, a 501c 3 organization).

*Ethical Statement:* The authors are accountable for all aspects of the work in ensuring that questions related to the accuracy or integrity of any part of the work are appropriately investigated and resolved.

*Disclaimer:* The views expressed in this editorial represent those of the authors and no other entity or organization.

**References**


