Pain Physician 2013; 16:E51-E59 • ISSN 2150-1149

Editorial Perspective

Spine Pain Classification: A Solution

Rinoo V. Shah, MD

Premise

A recent paper demonstrated that "magnetic resonance imaging (MRI) does not improve outcomes in patients who are clinical candidates for epidural steroid injections (ESI) and has only a minor effect on decision making"(1). This study has major clinical implications because it defies conventional wisdom: treatment is predicated upon appropriate patient selection by an etiological diagnosis. Imagine a patient undergoing an ESI by merely passing through easy diagnostic filters (clinical history, physical exam) and circumventing another (MRI). A physical exam is an easy diagnostic filter, since it has poor diagnostic accuracy for sciatica (2-5). If the minimum criterion is pain reproduction with a dural stretch or straight leg maneuver, how many patients would needlessly undergo an ESI? To ignore or not obtain a pre-procedure MRI suggests that patients without pathology could undergo an ESI. Yes, an MRI does have limitations in diagnosing radicular pain, but it would at least identify a potential structural etiology (6,7). Without an MRI, patients with normal spines could undergo an ESI. An MRI, furthermore, would nudge practitioners toward performing ESIs with target specificity, i.e., with image guidance. Paradoxically, ESIs double not only as treatment, but as indirect diagnostic tests; ESI success or failure informs future medical decisions. Should the ESI be repeated at a different level, should a different ESI technique be used, or should the patient be referred for other therapy? Nerve blocks are poor diagnostic tests (8), but represent well-intentioned, albeit archaic, efforts to improve patient selection. At a minimum, these diagnostic tests winnow the ESI candidate pool, but Cohen et al's (1) paper leads this author to conclude otherwise: an etiology-based diagnosis prior to treatment is not necessary (9).

This author previously demonstrated why spinal pain classification based on etiology is problematic and why alternatives should be sought (9). Spinal pain neurobiology is complex and constantly changing: a molecular black box. Since any etiological assumption about the events in this box inform treatment, accuracy is imperative. Unfortunately, most diagnoses are only guesses and patients and practitioners fail to recognize this. I propose an alternative classification system based on how patients fare with treatment. This is a post-hoc classification system, a treatment outcomes (TO) classification system.

Background

Despite guidelines, daily clinical practice is typically not evidence-based. Patients seek and practitioners deliver spine care as they see fit. Accountability, while preserving patient and practitioner autonomy, would reconcile this contest of wills. These are the cornerstones of the TO system.

Outcome Reporting and Framework

Patients would report their treatment outcome as a "tag" with 3 possible choices: before treatment (naught) or after treatment, success (S) or failure (F). This would be recorded at a specified time point, e.g., one week for chiropractic, 3 months for an epidural steroid, or one year for a laminectomy. Over a lifespan, the tags coalesce into a chain: a summary of a patient's treatment choices and outcomes.

A tag could be expanded or condensed to emphasize or minimize a treatment sphere. The chain would detail

From: Guthrie Clinic- Big Flats, Horseheads, NY.

Address Correspondence:
Rinoo V. Shah, MD, Guthrie Clinic-Big Flats, 31 Arnot Road,
Horseheads, NY 14845
E-mail: rinoo_shah@ahoo.com

Manuscript received: 09-18-2012
Revised manuscript received: 11-28-2012
Accepted for publication: 12-12-2012

changes accordingly. Imagine a patient chain with 3 tags: nonsurgical, surgical, and interventional. A pain physician could expand the interventional tag into 7 parts (facet injections, sacroiliac injections, epidural steroids, epidural lysis of adhesions, percutaneous disc decompression, radiofrequency neurolysis, and spinal cord stimulation). This expanded 9 tag chain is relevant to the pain physician, but the original 3 tag chain would interest a psychologist.

Tags could be represented by an ideogram, e.g., letters or images. In Fig. 1A individual tags are represented by circles. The fill represents outcomes: transparent means that treatment has not started, gray means treatment failure, and black means treatment success.

When individual tags are linked, they form a chain demonstrating the treatment choices and outcomes for a patient (Fig. 1B). In Fig. 1B, this patient has failed chiropractic care (ch), acupuncture (ac), physical therapy (pt), and non-opioid analgesics (no) as demonstrated by the first 4 gray circles; the patient succeeded with opioids (op) and surgery (su) as represented by the next 2 black circles; and finally, the patient did not undergo

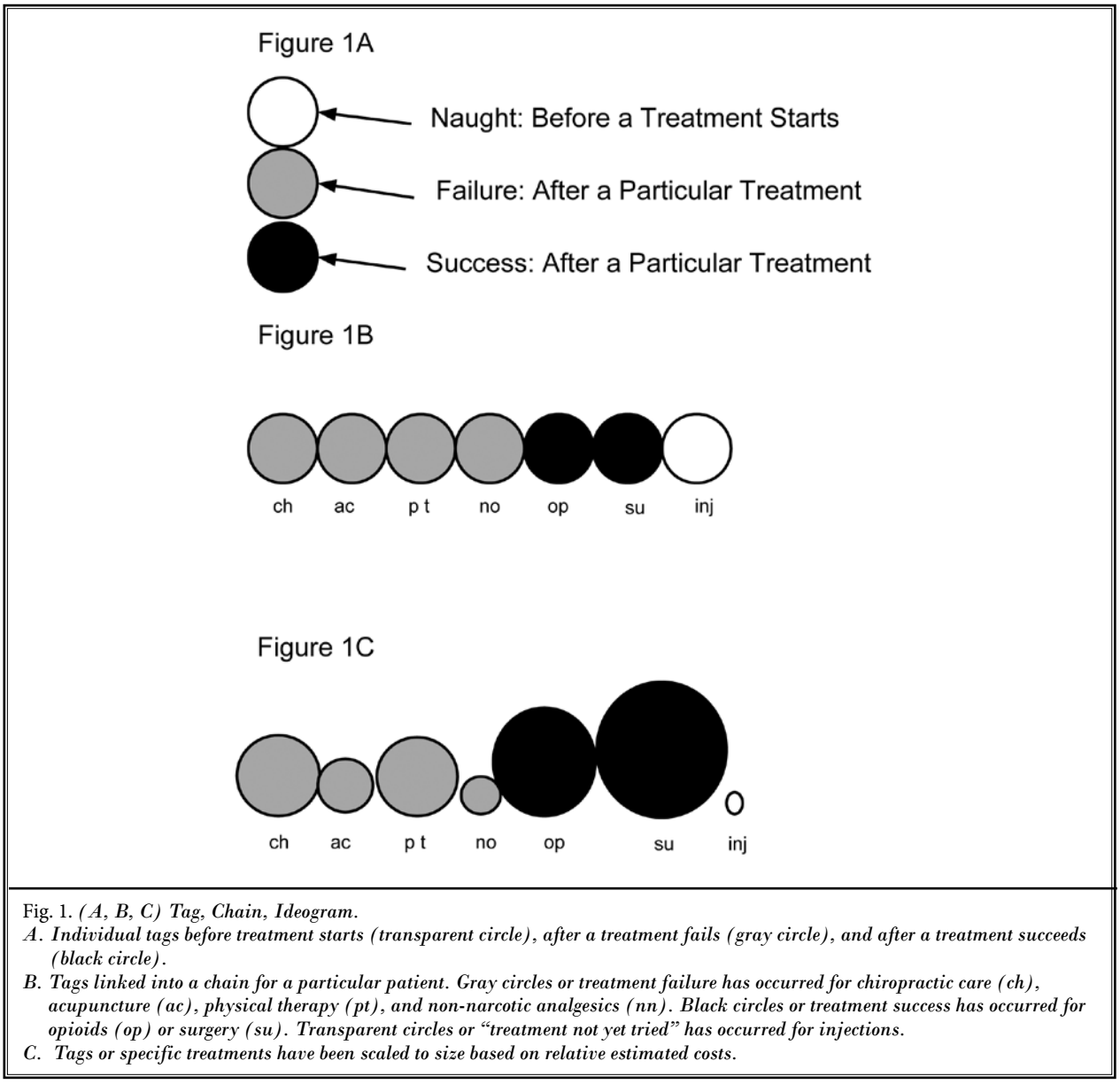

Fig. 1. *(A, B, C) Tag, Chain, Ideogram.*

A. Individual tags before treatment starts (transparent circle), after a treatment fails (gray circle), and after a treatment succeeds (black circle).

B. Tags linked into a chain for a particular patient. Gray circles or treatment failure has occurred for chiropractic care (ch), acupuncture (ac), physical therapy (pt), and non-narcotic analgesics (nn). Black circles or treatment success has occurred for opioids (op) or surgery (su). Transparent circles or "treatment not yet tried" has occurred for injections.

C. Tags or specific treatments have been scaled to size based on relative estimated costs.

injections (inj) as represented by the final transparent circle. The ideogram has been annotated to describe the specific treatments and their outcomes. Alternatively, circles and colors could be encoded by text FFFFSSø: F = failure; S = success; and ø = naught.

Visual Presentation of Data

Patient narratives and specialist evaluations are overly descriptive, fragmented, and interleaved with a discordant dialectic. The TO system distills complexity to a salient and exchangeable core. The image-based ideogram enables codification, outcome analysis between patients, and cost displays by scaling circle size (Fig. 1C). This ideogram, Fig. 1C, suggests that opioids and surgery are more effective, but more expensive than other treatments. Interventional procedures haven't been tried, hence, zero cost.

Unlike a patient narrative, diagnostic evaluation, and practitioner's assessment, the ideogram is an efficient communication tool for patients, practitioners, payers, and attorneys.

Patient Differentiation and Treatment Roadmap

The TO system is able to differentiate individual patients and guide future treatment. Two patients with similar diagnoses, e.g., degenerative disc disease, would be differentiated by personal treatment choices and outcomes. Repeat interventions and medications are often ordered unnecessarily. An ideogram would force all stakeholders to acknowledge a prior treatment outcome before repeating treatment. Decisions, without the burden of sending additional information and formal appeals, could be made quickly. An ideogram is transparent, preserves accountability, facilitates communication, and helps coordinate care.

Practitioner Benefits

An ideogram reduces chart/paperwork mining and associated data fatigue. A consult request could be: "40 year-old male with chronic low back pain that failed physical therapy, nonopioid analgesics, chiropractic care, and facet injections." A practitioner could make rapid, but informed decisions. Unlike an etiology-based diagnosis, a TO system diagnosis is easy to to comprehend and does not change arbitrarily

A practitioner could audit his own practice. For instance, a physician may notice that physical therapy treatment failures tend to fail facet injections. In a TO system analysis, this likelihood could be quantified and future patients counseled: "You failed prior physical therapy. My practice analysis suggests that similar patients are likely to fail facet injections. Let us explore alternatives."

Practices and payors select or reject patients for procedures, based on co-morbidities, e.g., poverty, obesity, smoking, or post-laminectomy syndrome. Instead of using these surrogate measures of future outcomes, actual historical spine care outcomes could be used! Denying treatment, for instance, because a patient doesn't have the 'right' insurance or is unemployed, is pejorative. Denying treatment, however, because of prior poor treatment outcomes, is not abusive; furthermore, this approach maintains patient accountability. Impeding patient access to treatment is an ethical issue, but it happens. A TO system analysis would introduce an 'arms-length' objectivity in treatment planning. Practitioners would still be able to shape their practice population, but in an ethical manner.

Informed Consent

Patients ask "I have heard that this procedure doesn't always work; what are the chances that it would work on me?" A response may be: "The literature suggests X% of patients have Y% pain relief, but I don't know how you will do."

The TO system enables use of an individual practice's data to complement the evidence-based literature. Cohort analyses of patient chains (outcome history) would estimate the likelihood of success versus failure for a future tag (procedure, medication, treatment). Now a practitioner could state: "Among patients that chose prior treatments identical to your choices, a subset had outcomes identical to yours. This subset consented to the procedure I wish to perform on you. These X patients had Y% success in my hands. My personal success rate in this patient population compares favorably to the evidence-based literature."

Informed consent in a TO system incorporates personalized outcomes and prognosis in addition to impersonal risks, benefits, and alternatives. Informed consent would no longer be an abstraction; it would be tailored to a particular doctor-patient-procedure nexus.

Reference Class Forecasting

Daniel Kahnemann (10), a Nobel Laureate in Economics, pioneered Reference Class Forecasting (RCF) in order to improve the reliability of forecasts. RCF takes an outside view versus the traditional inside view. The inside view is one that experts spontaneously adopt. There is a tight focus on the case at hand

(objectives, resources, obstacles), while extrapolating current progress into the future. Unfortunately, this approach is exceedingly optimistic. An etiology-based classification system uses the inside view. RCF uses an outside view, which ignores the details of the project at hand and future influences. RCF looks at similar projects and evaluates the outcomes of the routes taken by each project team. A probability distribution may then be created. The likelihood of the success of the current project can be viewed against this historical distribution. Lovallo and Kahneman (10) found this approach improves accuracy and reduces the influence of optimism, a psychological bias. The TO system lends itself to RCF analysis: a decision tree, with success/failure likelihood limbs, could be created. A patient and practitioner could use this tree for treatment planning.

Practitioner Risks and Accountability

Partition Coefficient

For a particular procedure, e.g., lumbar transforaminal, a practitioner will accumulate success-tags and failure-tags. As more patients undergo this procedure (increased sample size), a partition coefficient (Fig. 2) or ratio (success versus failure) can be generated. Ratios are recorded for each procedure performed by the practitioner. This is akin to a student receiving a report card for different school subjects.

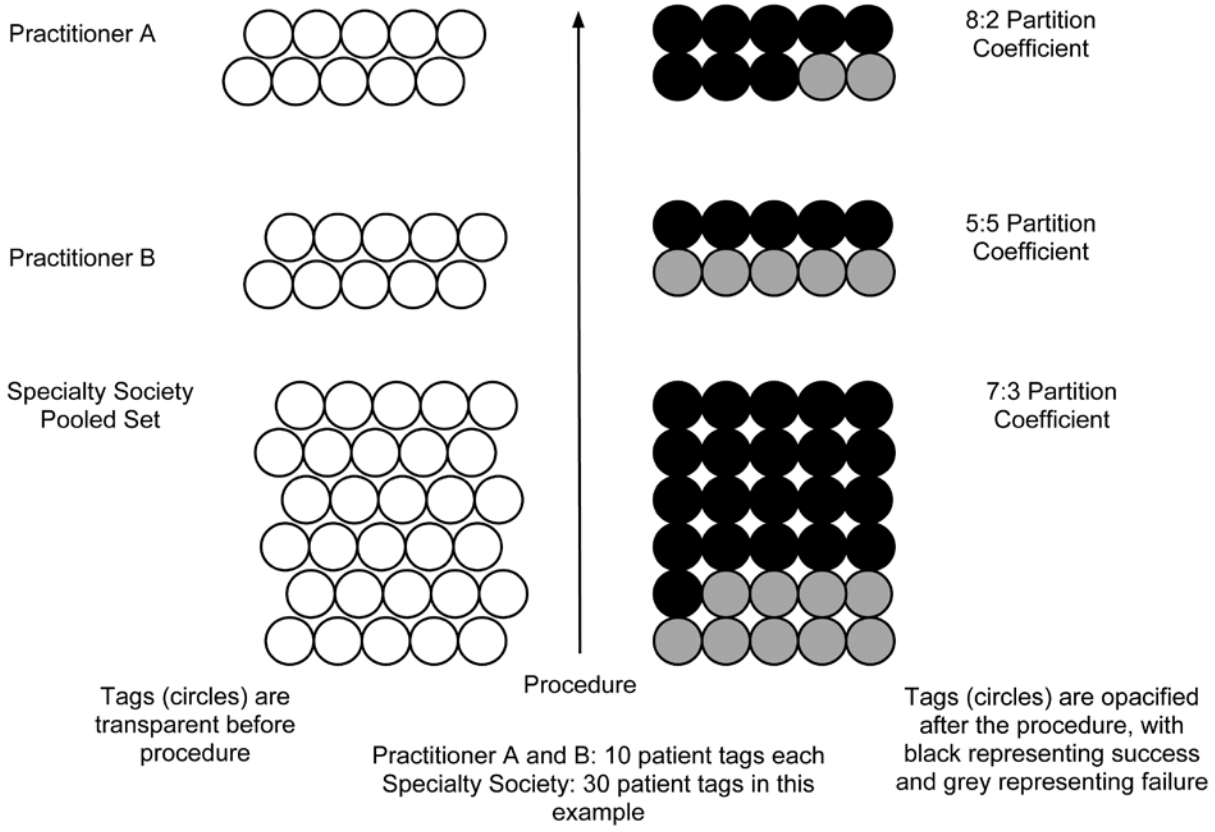

Fig. 2. *Provider vs. Specialty Partition Coefficients*
In this example, practitioners A and B have 10 patients each. The specialty society set contains 30 patients (arbitrary number for figure demonstration). The specialty society could maintain a registry for each procedure and the number of patients undergoing a particular procedure would vary. All of these patients undergo a procedure in the 'hands of' practitioner A, practitioner B, and 'all practitioners' in the specialty society. Each patient is represented as a transparent circle (tag), before the procedure. After procedure completion, the circles are opacified to grey (procedure failure) or black (procedure success). Practitioner A has 8 successes for 2 failures, i.e., an 8:2 partition coefficient or 80% success. Practitioner B has a 5:5 partition coefficient. The specialty society has a 7:3 partition coefficient. Practitioner A performs above par with respect to the specialty society, whereas Practitioner B is below par.

The evidence-based literature presents procedural outcomes as performed by experts on an optimal patient population. Partition coefficients would present procedural outcomes as performed by a practitioner and as reported by their patients. Evidence-based outcomes are vetted statistically and partition coefficients are not. However, a partition coefficient is a useful numerical heuristic; it is a rough assessment of the practitioner's patient selection and technical expertise for a procedure or treatment. Partition coefficients size up the real world implementation of spinal treatments by practitioners in daily practice. A TO system would make practitioners accountable.

This is analogous to learning a new subject. Early testing will demonstrate a wide variation in a student's exam scores; these should not be counted. With ongoing study, later exam scores will fluctuate to a lesser degree. Final exam scores will differentiate each student's mastery of a subject. Following a grace period (proximal edge of a learning curve), partition coefficient measuring should commence. With expertise (larger sample size), the partition coefficient should stabilize (tighter confidence interval, distal edge of a learning curve). Procedure or treatment mastery will yield partition coefficients that differentiate practitioners: different grades for different practitioners.

Decision Analysis and Risk Management

A TO framework facilitates decision-making and risk management. A practitioner would enter a patient chain (input) and different procedures (operators) into a database. With appropriate software, a distribution of procedure partition coefficients (success:failure) would be the outputs. Patient and practitioner could jointly view the outcomes of similar patients. Some procedures will have a higher partition coefficient than others. Patient and practitioner are free to choose any course of action, but all stakeholders would be informed of the likelihood of success.

If a surgeon decides to go against this information, the surgeon is free to do so. The patient would be informed ahead of time. A risk management company and payer could also weigh in. Surgeon and patient autonomy are preserved; however, control measures, i.e., financial incentives and accountability, are in place as well.

Audits and Investigations

An audit scrutinizes the intensity and frequency of a service by examining billing codes; patient outcomes are ignored. A physician that performs large numbers of epidurals or laminectomies could face penalties, even if their patients fared well. In a TO system, the focus shifts to patient outcomes; this feedback permits practice improvement, rather than just being punitive.

Partition coefficients would differentiate practitioners from peers. Subpar performance could be due to poor patient selection, poor technique, or difficult patients. The first 2 implicate the practitioner while the latter implicates the patient. A TO system would resolve debates about why the partition coefficient (practitioner grade) was poor.

Specialty societies could solicit members' partition coefficients and create a pooled set. While an auditor may rely on geographic- or community-based data, practitioners could use peer group data. Four scenarios are possible (Fig. 3). A partition coefficient allows everyone to understand a practice (Figs. 2, 3, 4), including medical boards.

In Fig. 2, partition coefficients are shown for practitioners and a specialty society. In Fig. 3, differences in outcomes between a practitioner and the specialty society could be investigated. Figure 4 looks at how physician selection of patients could influence a partition coefficient.

What if a high/low partition coefficient is because the practitioner selects easier or tougher patients? Consider a physician who has consistently favorable outcomes among patients that have either succeeded or failed prior therapies (Fig. 4, practitioner A); an isolated patient complaint could be viewed in the context of practitioner A's partition coefficient. Practitioner A has selected patients representative of his/her practice for treatment and still generates favorable outcomes, an 8:2 partition coefficient.

Conversely, a medical board may identify physicians that are a risk to the public via partition coefficients (Fig. 4, practitioner B). Practitioner B has chosen easier patients and denied difficult patients for treatment—not representative of his/her practice. Nonetheless, practitioner B still generates poor outcomes, a 4:6 partition coefficient. Practitioner A has not only selected a tougher patient population, but has performed better than practitioner B. A patient complaint or legal proceeding against practitioner A or B would now be viewed in context. In fact, experts deposed or invited to testify could be asked to produce their partition coefficients. The defendant practitioner could ask that selected experts meet a certain partition coefficient threshold. The TO system facilitates fairness and preemption in lieu of the current complaint-driven process.

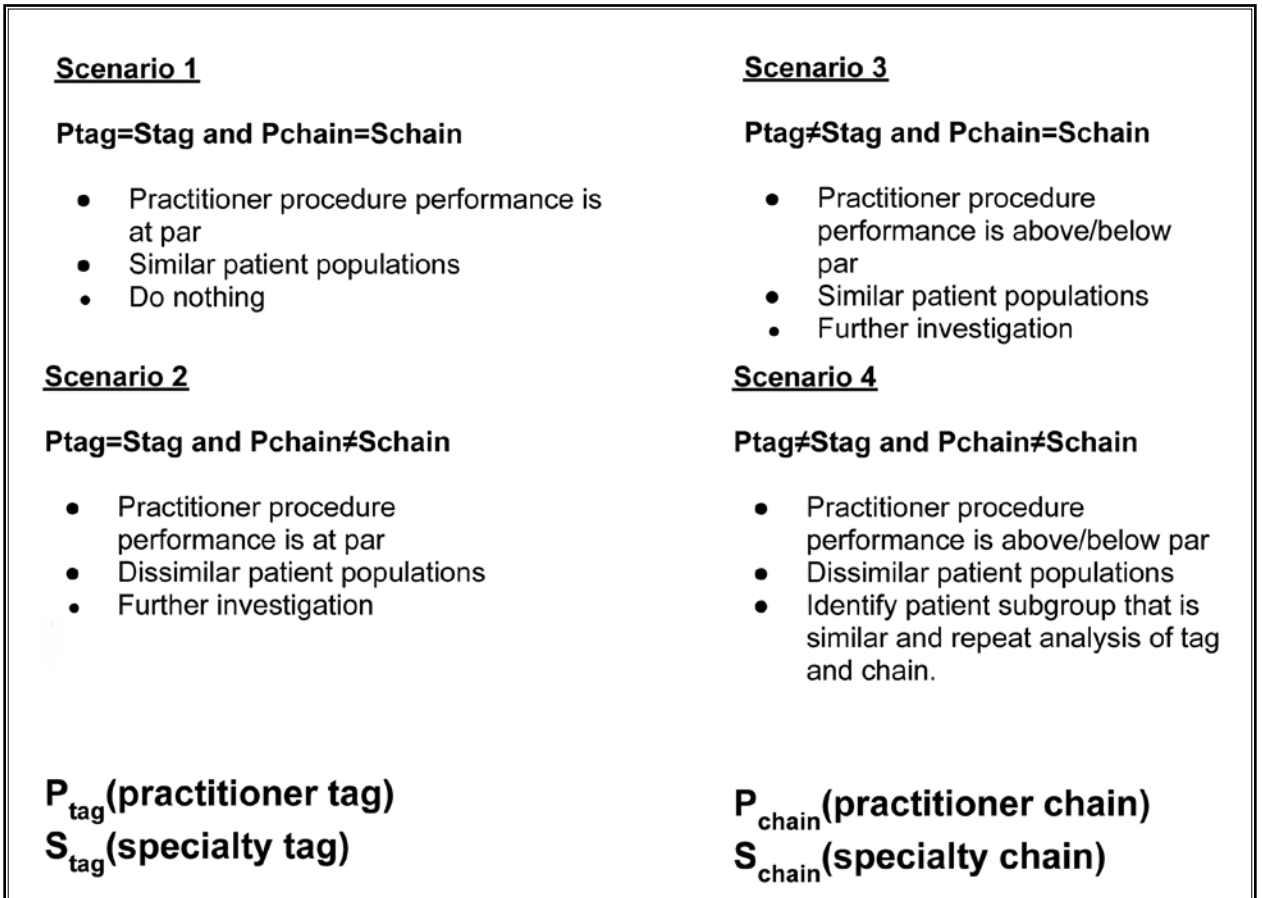

Fig. 3. Practitioner versus Specialty Society Comparisons.
Ptag is the pooled outcome of all patients undergoing a specific procedure, by a practitioner. Stag is the pooled outcome of all patients undergoing a specific procedure, by all practitioners within the same specialty. Pchain is the pooled outcomes of all patients in the practitioner's practice—his/her patient population as represented by their prior outcome histories. Schain is the pooled outcomes of all patients in a specialty society's practice (drawn from all members)—this collective patient population is represented by their prior outcome histories.
Four scenarios are possible.
Scenario 1: similar patient populations and outcomes between practitioner and specialty society;
Scenario 2: dissimilar patient populations, but outcomes are similar between practitioner and specialty society;
Scenario 3: similar patient populations, but outcomes are dissimilar between practitioner and specialty society;
Scenario 4: dissimilar patient populations and outcomes between practitioner and specialty society.
Conclusions for scenarios 1 – 4 are listed.

Research, Innovation, and Resource Allocation

Collective outcomes of an (index) procedure, such as all sacroiliac joint (SIJ) injections in 2010, are separated into groups: success versus failure. One then looks back at the linked tags and chains in each group. For instance, the SIJ success group may have an 80(s):20(f) chiropractic ratio; the SIJ failure group may have a 30(s):70(f) chiropractic ratio. Although suggestive, is there a relationship between chiropractic and SIJ outcomes? With wider TO adoption, a prospective analysis would replace this retrospective analysis. Then, predictive values could be calculated. This effectively joins completely different treatment spheres. Payers and practitioners could make patient-centered decisions, instead of just following blanket policies or society guidelines.

Payers are hesitant to pay for new technologies in the absence of high quality clinical evidence and label

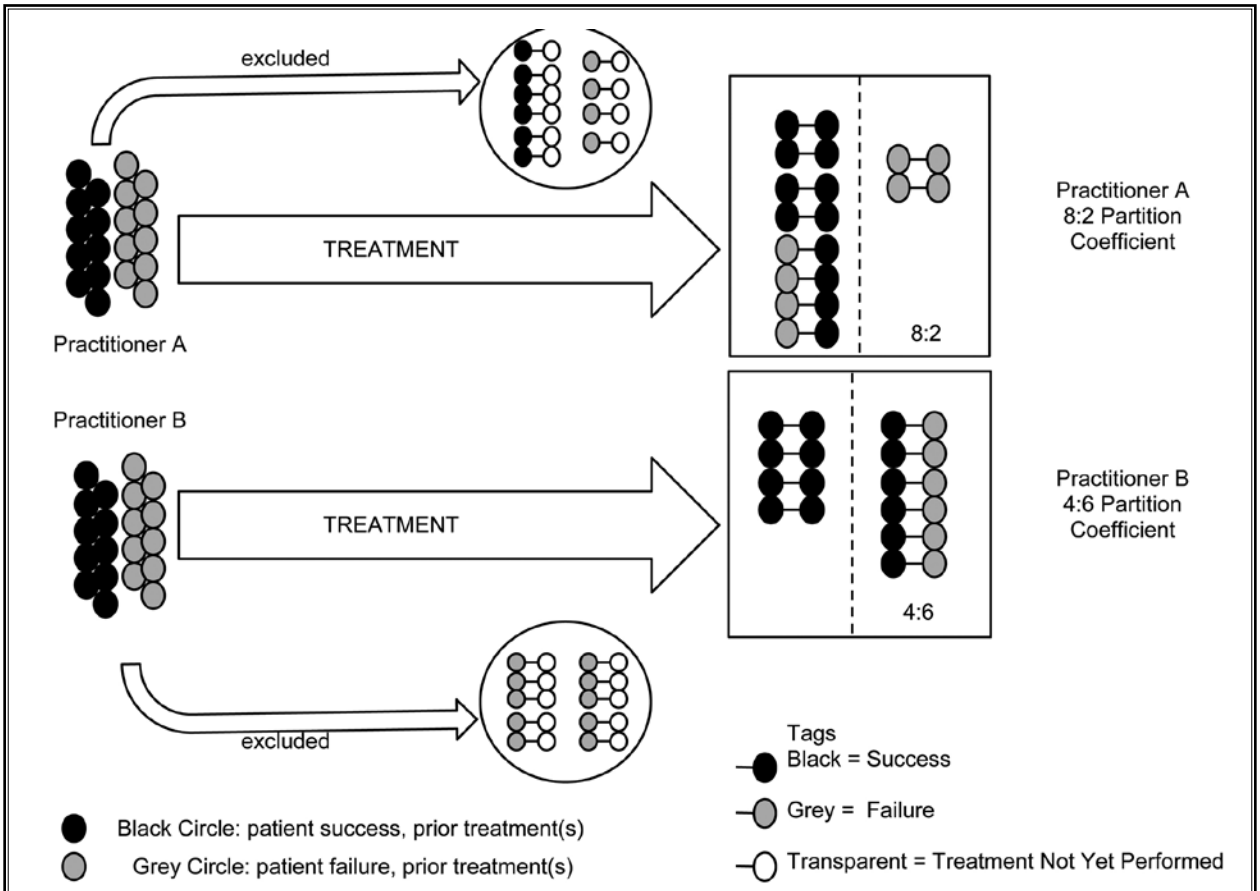

Fig. 4. *Practitioner Choice, Partition Coefficients, and Differentiation.*
Practitioners A and B have 20 patients each, represented by 20 circles. For each practitioner, 10 patients have failed prior treatments (10 gray circles) and 10 patients have succeeded with prior treatments (10 black circles). Practitioners A and B have excluded 10 patients each from getting treatment; these black or gray circles are now linked to transparent circles—indicating that they did not get this treatment. Practitioner A has excluded 6 prior outcome success (black circles) and 4 prior outcome failure (gray circles) patients. Practitioner B "cherry picked," by excluding all 10 patients with prior outcome failures (gray circles). Practitioner A has selected a tougher population of patients than practitioner B for treatment. These selected patients then report their outcomes for this new treatment. Practitioner A has 8 patients succeeding and 2 failing or an 8:2 partition coefficient. Practitioner B has 4 patients succeeding and 6 failing or a 4:6 partition coefficient. In other words, practitioner B has cherry picked patients but still has poorer outcomes than practitioner A.

many treatments as investigational. This has led to a surfeit of direct to consumer marketing for procedures, e.g., minimally invasive spine surgery and nonsurgical spinal decompression. The TO system would allow new treatments to compete and would ascertain the procedure partition coefficient (success:failure). Payment could be structured as a bonus or penalty, depending on the outcomes of the technique. Payers would be able to play an active role in innovation and patient care in addition to the one they play in cost containment.

The TO system will identify the provider, patient, and procedure combination that yields the best outcome. Patients could be nudged toward a combination by using payment as an incentive or disincentive. Accountability and direct costs can be traced to the patient, practitioner, and treatment. Resources would be allocated based on outcomes.

Limitations

This is a post-hoc classification based on patient self-reporting. Patient reporting of outcomes may be heavily influenced by perceptual factors, challenges in communication, or external irrelevant information. For instance, patients often misunderstand the

term pulmonary nodule as cancer (11). Practitioner communication can heavily influence patient perception. . Furthermore, advocacy by patient groups with marketing may heavily influence patient opinion (12). Nonetheless, using patient rating instruments, e.g., the Oswestry Disability Index or numeric rating scales, is commonplace in pain management research (13,14). This information is solicited in daily clinical practice, especially hospitals. Self-reporting rating instruments have reliability in detecting improvement or worsening in most patients (13). Due to this, clinical use is advised (13,14). However, routine outpatient use of self-report scales is uncommon. These reporting instruments take time and require the assistance of a trained individual. The TO system simplifies feedback and would be easily implemented in a clinic.

The purpose of soliciting patient feedback is to improve patient care and to audit the success or failure of a particular pain treatment. Recently, Web sites such as Patients Like Me have succeeded with crowd-sourcing patient health information reporting (15). Patients can freely report their findings; this may act to support or to contradict the results of a clinical trial. The purpose of the Patients Like Me website is to help 'people with every type of condition are coming together to share their health experiences, find patients like them and learn how to take control of their health. The result is improved care for patients as well as an acceleration of real-world medical research'. Crowd-sourcing, however, is vulnerable to cacophony, since individuals providing their input could be anonymous and unaccountable. The TO system, on the other hand, preserves accountability and lends itself to quantitative analysis.

Another criticism is that the TO system appears to simplify outcomes. Practitioners may feel angered to have their patient care efforts reduced to success or failure. The current model is "first generation." Binary reporting could eventually morph into a spectrum. Patients could report a range of outcomes from 0% (absolute failure) to 100% (absolute success). This author believes this is not necessary. Patients may have difficulty reporting percentage outcomes, particularly individuals with poor math skills. Any patient reporting 'tool' should be simple. Furthermore, practitioners outside the specialty may not desire nuanced outcomes. Is an interventional pain physician interested in a 50% physical therapy success? Is a chiropractor interested in a 20% surgical success? A binary system of success versus failure should suffice.

Conclusion

A Treatment Outcomes classification system for spine pain is proposed. This system could run in parallel and eventually, replace the etiology based spine pain classification system. Overall, this Treatment Outcomes system would preserve autonomy, improve accountability, facilitate innovation and research, improve fairness, foster competition, shift focus to resource allocation, promote education, and improve the informed consent process. The Treatment Outcomes system would level the playing field for all treatment approaches in spine care, without concerning itself with the 'black box' of spine pain etiology.

References

1. Cohen SP, Gupta A, Strassels SA, Christo PJ, Erdek MA, Griffith SR, Kurihara C, Buckenmaier CC 3rd, Cornblath D, Vu TN. Effect of MRI on treatment results or decision making in patients with lumbosacral radiculopathy referred for epidural steroid injections: A multicenter, randomized controlled trial. *Arch Intern Med* 2012; 172:134-142.
2. Gilbert KK, Brismée JM, Collins DL, James CR, Shah RV, Sawyer SF, Sizer PS Jr. 2006 Young Investigator Award Winner: Lumbosacral nerve root displacement and strain: Part 2. A comparison of 2 straight leg raise conditions in unembalmed cadavers. *Spine (Phila Pa 1976)* 2007; 32:1521-1525.
3. Gilbert KK, Brismée JM, Collins DL, James CR, Shah RV, Sawyer SF, Sizer PS Jr. 2006 Young Investigator Award Winner: Lumbosacral nerve root displacement and strain: Part 1. A novel measurement technique during straight leg raise in unembalmed cadavers. *Spine (Phila Pa 1976)* 2007; 32:1513-1520.
4. Devillé WL, van der Windt DA, Dzaferagić A, Bezemer PD, Bouter LM. The test of Lasègue: Systematic review of the accuracy in diagnosing herniated discs. *Spine (Phila Pa 1976)* 2000; 25:1140-1147.
5. Hancock MJ, Koes B, Ostelo R, Peul W. Diagnostic accuracy of the clinical examination in identifying the level of herniation in patients with sciatica. *Spine (Phila Pa 1976)* 2011; 36:E712-E719.
6. Boos N, Rieder R, Schade V, Spratt KF, Semmer N, Aebi M. 1995 Volvo Award in clinical sciences. The diagnostic accuracy of magnetic resonance imaging, work perception, and psychosocial factors in identifying symptomatic disc herniations. *Spine (Phila Pa 1976)* 1995; 20:2613-2625.
7. Wassenaar M, van Rijn RM, van Tulder MW, Verhagen AP, van der Windt DA, Koes BW, de Boer MR, Ginai AZ, Ostelo RW. Magnetic resonance imaging for diagnosing lumbar spinal pathology in adult patients with low back pain or sciatica: A diagnostic systematic review. *Eur Spine J* 2012; 21:220-227.
8. Shah RV. The problem with diagnostic selective nerve root blocks. *Spine (Phila Pa 1976)* 2012; 37:1991-1993.

9. Shah RV. Spine pain classification: The problem. *Spine (Phila Pa 1976)* 2012; 37:1853-1855
10. Lovallo, D., Kahneman, D. Delusions of success: How optimism undermines executives' decisions. *Harvard Business Review* 81, July 2003. http://hbswk.hbs.edu/archive/3630.html
11. Wiener RS, Gould MK, Woloshin S, Schwartz LM, Clark JA. "What do you mean, a spot?": A qualitative analysis of patients' reactions to discussions with their doctors about pulmonary nodules. *Chest* 2012; Epub ahead of print.
12. Woloshin S, Schwartz LM. How a charity oversells mammography. *BMJ* 2012; 345:e5132.
13. Davidson M, Keating JL. A comparison of five low back disability questionnaires: Reliability and responsiveness. *Phys Ther* 2002; 82:8-24.
14. DeVine J, Norvell DC, Ecker E, Fourney DR, Vaccaro A, Wang J, Andersson G. Evaluating the correlation and responsiveness of patient-reported pain with function and quality-of-life outcomes after spine surgery. *Spine (Phila Pa 1976)* 2011; 36:S69-74.
15. Patients Like Me. www.patientslikeme.com