

Letters to the Editor-in-Chief

Researchers and Clinicians: A Growing Divide or Narrowing Gap?

J Orthop Sports Phys Ther 2006;36(12):963. doi:10.2519/jospt.2006.0202

I would like to comment on the guest editorial, “Researchers and Clinicians: A Growing Divide or Narrowing Gap?”¹ by Roy Bechtel, Joshua A. Cleland, and Britt Smith, published in the July 2006 issue of the *JOSPT*. As a clinician and researcher for nearly 30 years, I find the discussion between the clinical and the scientific fascinating. The debate over empirical or practical versus the ultimate truth goes back at least to the fourth century before Christ and was famously depicted in a fresco by the artist Raphael. In this work, Raphael shows Aristotle gesturing to the earth, representing his belief in knowledge through empirical observation and experience (read clinical), while Plato points up to the heavens, reflecting his belief in the ultimate truth (read science). I believe this longstanding difference in approach should not be a shibboleth between clinician and researcher, but rather a driving force that keeps us moving forward in the search for greater knowledge and understanding. This, to me, is what we live for: a never-ending quest for the truth!

Physical therapists treat people with movement disorders. Thus, we mainly need clinical articles that focus on the diagnosis, prognosis, and intervention of movement disorders. Good science is a necessary, but not a sufficient requirement for a good clinical article. What is necessary and sufficient is clinical relevance. The modern father of clinical epidemiology, Alvin Feinstein, MD stated, “No research project can be perfectly conceived, conducted, or analyzed; imperfections can easily be found if enough items are checked.”² I would much rather read a clinically relevant article that has some methodological problems than one with no clinical relevance and few or even no methodological problems. The important point is that, inherently, most empirical or practical studies—no matter how good they are—will always have some type of scientific flaw. “The crucial issue is to decide which imperfections are important and which are not.”²

Ultimately, finding the truth requires a preponderance of evidence. The evidence should be from different researchers, methods, or types of research, and over different periods of time. The evidence should also include both convergent and divergent methods of establishing the proof. How to find the clinically important truth is the seminal question. The only way I see this happening is through the collaborative efforts of experienced clinicians and

researchers. It is no accident that the authors who publish great clinical articles that have the greatest impact on our profession have already figured out this secret.

Michael T. Cibulka, PT, DPT, MHS, OCS
Assistant Professor
Physical Therapy Program
Maryville University
St Louis, MO

1. Bechtel R, Cleland JA, Smith B. Researchers and Clinicians: A Growing Divide or Narrowing Gap? *J Orthop Sports Phys Ther*. 2006;36(7):451-461.
2. Feinstein, A. *Clinical Epidemiology. The Architecture of Clinical Research*. Philadelphia, PA: W.B. Saunders Company; 1985.

Patient Outcomes and Clinical Performance: Parallel Paths or Inextricable Links?

J Orthop Sports Phys Ther 2006;36(12):963-965. doi:10.2519/jospt.2006.0203

I would like the opportunity to express some concerns about the call for basing reimbursement for physical therapy services on a pay-for-performance (P4P) scheme and assessing student clinical performance based on patient outcomes, as proposed by Anthony Delitto, PT, PhD, FAPTA in his guest editorial in the August 2006 issue of the *JOSPT* (“Patient Outcomes and Clinical Performance: Parallel Paths or Inextricable Links?”).

While I fully share Dr Delitto’s frustration that too many clinicians are not taking into consideration the best available evidence when designing treatment plans and agree wholeheartedly that positive patient outcomes are the most important aspect of the therapeutic intervention, pay-for-performance and holding students up to some established percentage of positive outcome measurement are not the answers. Under pay-for-performance there would be an established number of visits in which a clinician would be expected to get patients with a given diagnosis better. If on average the clinician got those patients better in fewer visits, there would be an extra financial reward for the clinician to make up for the visits not used. If, on average, the clinician got those patients better in about the expected time, there would be no reward or perhaps some small reward. If, on average, it took the clinician longer to achieve the desired outcomes, the clinician would ultimately get less per visit than those who achieved those outcomes in less time. The descriptions of the pay-for-performance programs are phrased to make it sound like there are rewards for outstanding perfor-

mance without penalty for lesser performance; but, clearly, those who on average do not achieve the desired outcomes in a shorter period of time will be penalized. There is just so much money in the pot and, if some are given more as a reward, others will get less as a penalty.

On paper, as a basic concept, pay-for-performance sounds great. Clinicians should be paid based on performance, on achieving positive outcomes in a cost-effective manner. How can you argue against that? Unfortunately, what looks good in concept isn't always so good in actuality. Remember that managed care, capitation rates, and case rates also sounded like pretty good concepts. None of those worked very well. The problem with pay-for-performance is that the established expectations for the time it should take to achieve positive outcomes for a given diagnosis will be based on data collected from a large cohort of thousands of patients with that diagnosis. While it might be nice to have statistical measurements on large groups of people for research and actuarial purposes, those statistical measurements cannot be used to predict how a particular individual should or would respond to treatment in recovering from a disease or injury.

As the literature tells us, to date, we have no validated prognostic indicators, particularly in orthopedics. It is therefore impossible to prospectively identify outliers, those patients who will not respond in the expected time. We know that, even with the best of care, for acute musculoskeletal injuries somewhere between 2% and 15% of patients will go on to develop chronic pain.⁵ Even with the clinical prediction rule for classifying patients with low back pain who will respond to manipulation,^{3,4} we can only predict what is likely to happen with a small percentage of the many patients who will have an ICD-9 number for a diagnosis of low back pain. That would be only for those patients with low back pain who meet the criteria of the clinical prediction rule. The ICD-9 numbers that clinicians choose for low back pain are fairly general for low back pain such as 724.2 (Lumbago) or 847.2 (Sprain/Strain Lumbar). The diagnosis code will not tease out just those patients who have a low back condition that should respond to manipulation. They will all be lumped together with the rest of the patients with low back pain. Even of those who meet the criteria, not all will respond well to the chosen intervention, manipulation. The rule tells us that it would be wise to try treating those patients with manipulation, because they are likely to have a good outcome. But it certainly does not guarantee that all of those patients will have a positive timely outcome.

When measuring thousands of patients' outcomes, the outliers become lost in the averages and therefore do not become a problem. The problem occurs when the statistics are used to judge the quality of

care provided by an individual practitioner or even a small practice with just a few practitioners. If an individual sees 250 patients a year, he might see less than 20 patients a year with any particular diagnosis of, let's say, sprained knee. If just 1 of those patients turns out to be an outlier, legitimately requiring 30 visits rather than the expected 6, that therapist's statistics will be thrown far out of proportion to the expected norm and that practitioner, or small practice, will be unfairly penalized in a pay-for-performance system.

The percentage of outliers doesn't even have to be very large to have an adverse effect on a significant number of small practices. Let's assume that for a particular diagnosis there might only be 7 outliers for every 100 patients. For a large provider treating all of those 100 patients the 7 outliers might not hurt because they will be averaged out over the year. (Actually even 100 patients might not be a large enough number to offset the outliers; but for the sake of simplicity, I am using the 100 and 7 numbers for this example.) If, on the other hand, those 100 patients are treated by 10 different therapists, with each therapist seeing 10 patients, the advantages of averaging disappear. While it is unlikely that any one of the therapists will get stuck with all of the 7 outliers, it is equally unlikely that they will be distributed evenly so that no therapist gets more than 1.

If it is expected that the total number of treated patients with this diagnosis will, on average, get better in 6 visits and you have just 1 of the outliers in your 10, with no patients requiring the less than expected number of visits to set them off, then it is likely that you will not average 6 visits per patient. If 9 of the patients do average the 6 visits but the 10th patient takes 30 visits, the average number of visits to get those patients better at that clinic would be 8.4 and that clinic would be penalized. If you are asked to treat 2 or more of the outliers the numbers will be even worse. All of the physical therapists who treat the outliers will be performing below expectation and therefore subject to some penalty in a pay-for-performance scheme. If the outliers get distributed evenly, that means 7 of the 10 physical therapists will be sub par. If the outliers are distributed randomly, as they probably will be, then anywhere from 2 to 6 of the 10 physical therapists will be sub-par. They will all be penalized, while the remaining physical therapists, who simply by luck of the draw ended up with patients with good prognosis, will get rewarded, sometimes unjustly. That is not a very good scenario.

The American Medical Association (AMA) House of Delegates has adopted "Principles and Guidelines for Pay-for-Performance"¹ and made a recommendation "that our American Medical Association oppose private payer, Congressional, or Centers for Medicare and Medicaid Services pay-for-performance initiatives

if they do not meet the AMA's 'Principles and Guidelines for Pay-for-Performance.'" In a press release announcing the House action, an AMA Board of Trustees member acknowledged that current legislation before the Senate Finance Committee does not meet the AMA's principles and guidelines.

The APTA Board of Directors has also adopted a similar position on pay-for-performance to assure that, if it is ever implemented, it is done fairly.² No government or payer pay-for-performance proposals outlined to date would satisfy either AMA or APTA's positions. It is likely that given the basic flaws in the concept, as outlined above, they never will. For those reasons we should be very careful not to embrace or encourage pay-for-performance as a solution to clinical problems just yet, if ever. It is just not a good bandwagon for us to jump on.

The same statistical problems arise when we try to judge physical therapy students based on expected outcomes. Even though internships are now longer than they used to be, it is unlikely that a student will see enough patients with any one diagnosis, during any one internship, to overcome providing care to 1 or 2 patients who don't respond as expected. Conversely, a student might be fortunate enough during a particular internship to see only those patients who get better no matter how they are treated. We do know that some of those patients exist. Would we really want students' clinical performance rating to be based on those outcomes?

It has long been said that reimbursement should not drive practice. Practitioners should not use more interventions just to get more reimbursement, nor should they use fewer interventions or visits to do better under a capitation rate. Each patient should get what is needed, no more no less. Asking for pay-for-performance is asking for reimbursement to drive practice and that is not a good practice.

So yes, we do have to get all physical therapists on track to using the best evidence available to select

treatment options and to monitor outcomes of those treatments to assure that patients respond accordingly. But we cannot do that using external punitive forces. We have to do it the old fashion way, with education, peer pressure, and public awareness raising patients' expectations. Dr Delitto and many others are going a long way to make that happen with their outstanding research, teaching, and clinical accomplishments. It is certainly frustrating that everyone has not yet grasped and implemented these concepts and it would be nice to be able to speed the process. But that is just not the way social change works. The ultimate goal will take time to achieve. Most good things do.

Philip Paul Tygiel, PT, MTC
Tygiel Physical Therapy
Tucson, AZ

REFERENCES

1. American Medical Association. Board of Trustees Report 5 - Pay-for-Performance Principles and Guidelines. Available at: <http://www.ama-assn.org/meetings/public/annual05/bot5a05fin.doc>.
2. American Physical Therapy Association. Board of Director's Position. Available at: <http://www.apta.org/AM/Template.cfm?Section = Minutes1&Template = /MembersOnly.cfm&ContentID = 27481>.
3. Childs JD, Fritz JM, Flynn TW, et al. A clinical prediction rule to identify patients with low back pain most likely to benefit from spinal manipulation: a validation study. *Ann Intern Med.* 2004;141: 920-928.
4. Flynn TW, Fritz J, Whitman J, Wainner R, Magel J, Rendiero D, Butler B, Garber M, Allison S. A clinical prediction rule for classifying patients with low back pain who demonstrate short-term improvement with Spinal Manipulation. *Spine.* 2002;27:2835-2843.
5. Linton SJ, Helsing A, Anderson D. A controlled study of the effects of early intervention on acute musculoskeletal pain problems. *Pain.* 54 (1993) 353-359.

Authors' Response

J Orthop Sports Phys Ther 2006;36(12):965-967. doi:10.2519/jospt.2006.0203

It would appear that Mr Tygiel's definition of an outlier centers on a particular patient that does not improve during a physical therapy episode of care. Some of these "outliers" may legitimately require more intensive physical therapy care, as Mr Tygiel implies, while others may not improve at all with physical therapy intervention, a scenario that Mr Tygiel ignores. Mr Tygiel's fear is that treatment for individuals who require more intensive care will be curtailed by overgeneralizing information from the

literature, which defies a basic premise of evidence-based practice. It would be inappropriate and indefensible to generalize a prediction rule for manipulation developed for patients with acute low back pain to a population of people with chronic low back pain, for example, with the expectation of a similar degree of outcome changes in both groups. I do not think that anyone who supports an evidence-based approach to care is suggesting generalizing the literature in this manner. Mr Tygiel's challenges also ignore published evidence that render his whole outlier scenario baseless. For example, he makes the false statement that "we have no validated prognostic

indicators” and, therefore, it is “impossible to prospectively identify outliers.” We indeed have positive as well as negative prognostic indicators in musculoskeletal as well as other forms of physical therapy that have a very solid evidence basis.^{1-3,6-10,12} For example, there is some very solid evidence that fear-avoidance beliefs are prognostic for predicting patient’s ability to return to work at full duty.^{4,6} The challenge, as I see it, is to find a way to get clinicians to utilize this information in their everyday practice.

I would now like to focus my response on the patients who are not outliers and who fit the clinical prediction rule for manipulation perfectly—a scenario that can comprise a significant proportion of patients with acute low back pain. Mr Tygiel’s solution to the problem of inconsistency between clinician behavior and published evidence can be summarized by the statement “it will take time.” This “solution” is indefensible from the standpoint of its consequences (eg, continued suboptimal care in cases in which we know better can be provided) and its low likelihood of success. Physical therapists who purport themselves to be evidence-based practitioners and who fail to follow evidence in their clinical decision making are not only doing a disservice to their patients, but they are detrimental to a profession that bills itself to be evidence based. I would see some sort of penalization to be the quickest form of behavior change. Quite a while back, when I was a student, one of the first things I learned in the clinic was that even suggesting “maintenance” in a Medicare note was tantamount to denial. Therapists “learned” that writing such a word or any derivative of it would quickly lead to denial of the claim. Denial led to people never using that word in a Medicare note.

We often get lost in the details of measuring quality and best care, and developing good prognostic indicators. Yet we tend to forget the inertia that will need to be overcome by professional organizations (including our own APTA) that will tend to protect their membership’s interests against any scenario that resembles a loss. Consequently, it is highly unlikely that professional organizations will ever be inclined to advocate for reimbursement scenarios that result in a penalty. Couple this with every professional organization wanting added incentives for good performance and it becomes quickly clear that pay-for-performance (P4P) stipulations put forth by professional organizations (including the AMA, as outlined by Mr Tygiel) will result in added costs. This is a very shortsighted view given that the problem trying to be addressed with P4P is better utilization of existing resources.

Imagine a member of Congress sitting on committees anticipating hearing the details on how a P4P initiative will result in increased efficiencies only to be overwhelmed by testimony from professional groups protecting their respective turf by demanding no penalties yet asking for premiums based on good

performance. The end result will be the status quo, with no check whatsoever on inefficient care. Not even low-hanging fruit, such as acute low back pain! Therefore, we should not be counting on the government (eg, CMS) or our professional organization for “change agent” leadership to guide us out of the current healthcare dilemma. In fact, any “new P4P” initiative on their part is years away, because it’s a whole lot easier to not move than to move (the definition of inertia). It is more likely that employer groups, whose incentives are better aligned for the long run, will ultimately drive the change.

I would like to offer an alternative explanation to what I see as an overall hesitancy to endorse P4P initiatives. The real challenge lies in the fact that P4P requires a responsible and trusted partnership with the payer industry, many of whom have a long history of not being able to be trusted. I believe what lies behind the apparent obstinacy of professional organizations and individuals about considering rewards and penalties is mistrust of the payer industry, particularly when it comes to truly rewarding good performance. What many fear is the payer’s interpretation of “reward” to mean the status quo with regard to good performance (perform well and you get paid just like normal); perform poorly, however, and you will be penalized. This scenario obviously saves the payer money and adds to their already burgeoning coffers, but becomes a disincentive to the provider in that they have nothing to gain and potentially a lot to lose. Mistrust of the payer industry may indeed be one good reason for lobbying to keep the status quo in reimbursement guidelines for both physical therapy and medicine resulting in resistance to P4P initiatives. I can readily accept reasons centering on this mistrust rather than the limitations of evidence-based practice principles as espoused by Mr Tygiel.

Anthony Delitto, PT, PhD, FAPTA
Department of Physical Therapy
School of Health and Rehabilitation Sciences
University of Pittsburgh
Pittsburgh, PA

REFERENCES

1. Croft PR, Dunn KM, Raspe H. Course and prognosis of back pain in primary care: the epidemiological perspective.[comment]. *Pain*. 2006;122:1-3.
2. Dunn KM, Croft PR. The importance of symptom duration in determining prognosis. *Pain*. 2006;121:126-132.
3. Faber E, Burdorf A, Bierma-Zeinstra SM, Miedema HS, Koes BW. Determinants for improvement in different back pain measures and their influence on the duration of sickness absence. *Spine*. 2006;31:1477-1483.
4. George SZ, Bialosky JE, Donald DA. The centralization phenomenon and fear-avoidance beliefs as prognostic

- factors for acute low back pain: a preliminary investigation involving patients classified for specific exercise. *J Orthop Sports Phys Ther.* 2005;35:580-588.
5. George SZ, Bialosky JE, Fritz JM. Physical therapist management of a patient with acute low back pain and elevated fear-avoidance beliefs. *Phys Ther.* 2004;84:538-549.
 6. George SZ, Fritz JM, Bialosky JE, Donald DA. The effect of a fear-avoidance-based physical therapy intervention for patients with acute low back pain: results of a randomized clinical trial. *Spine.* 2003;28:2551-2560.
 7. Jones GT, Johnson RE, Wiles NJ, Chaddock C, Potter RG, Roberts C et al. Predicting persistent disabling low back pain in general practice: a prospective cohort study.[see comment]. *Brit J Gen Pract.* 2006;56:334-341.
 8. Lotters F, Burdorf A. Prognostic factors for duration of sickness absence due to musculoskeletal disorders. *Clin J Pain.* 2006;22:212-21.
 9. Pransky GS, Verma SK, Okurowski L, Webster B. Length of disability prognosis in acute occupational low back pain: development and testing of a practical approach. *Spine.* 2006;31:690-697.
 10. Schultz IZ, Crook J, Berkowitz J, Milner R, Meloche GR. Predicting return to work after low back injury using the Psychosocial Risk for Occupational Disability Instrument: a validation study. *J Occup Rehab.* 2005;15:365-376.
 11. Shaw WS, Pransky G, Patterson W, Winters T. Early disability risk factors for low back pain assessed at outpatient occupational health clinics. *Spine.* 2005;30:572-580.
 12. Steenstra IA, Verbeek JH, Heymans MW, Bongers PM. Prognostic factors for duration of sick leave in patients sick listed with acute low back pain: a systematic review of the literature [review]. *Occup Environ Med.* 2005;62:851-60, 2005.