

The Q & A on P4P

Is pay-for-performance new?

No. Private health plans have been utilizing various formats of pay-for-performance for several years. Almost all of the programs focus exclusively on two areas of care – primary care and cardiology because those are the two areas with well-accepted, evidence-based quality measures. In addition, cardiac thoracic surgery has been involved in several demonstration projects utilizing its cardiac surgery outcomes database.

What Lessons Have Private Payers Learned About Pay-for-Performance?

Some private pay payer programs have been successful while others have not. Successful programs have several things in common:

1. Physician reimbursement from the health plan was stable and predictable before implementation of pay-for-performance;
2. They involve the investment of “new money;” and incentive payments made to top performers are truly bonus payments;
3. Incentive payments must be a meaningful amount, predictable and timely;
4. Quality measures used were well-accepted and evidence-based;
5. Actual improvement in patient care was attainable and measurable

Can All Specialties Participate in CMS’ Physician Voluntary Reporting Program?

No. Many specialties, including those involving significant Medicare volume like cataract surgery and dermatology, cannot participate because there are no available, fully-approved quality measures for these areas. Almost all of the work taken on by groups like the AMA Physician Consortium and the National Quality Forum involve primary care and cardiology.

Why Aren’t There Quality Measures for Some Specialties?

There are a variety of reasons why the majority of measures available are in primary care and cardiology. One of the most common problems revolves around the availability of clinical evidence. In order to develop quality measures, medical societies must first develop clinical practice guidelines. In order to develop guidelines, there has to be evidence to support the “evidence base.” Quality clinical guidelines are developed from Level I evidence, which is defined as evidence derived from randomized, controlled clinical trials. It is impossible for the surgical specialties to obtain this level of evidence because randomized, controlled studies cannot be performed in surgery (unlike a drug therapy where placebo drugs can be tested, performing “placebo surgery” – essentially putting an patient under anesthesia, opening them up but not actually surgically addressing the problem – is controversial with few willing participants). In addition, many types of Level II evidence are also difficult to obtain for many specialties for the same reason.

In addition to the logistical problems involving generating evidence for many specialties, there are also financial constraints. Because the prevalence and incidence of many diseases treated by specialists are small compared to heart disease, diabetes and the entire realm of primary care, research funding is not nearly as great. This is also the reason why many of the guidelines specialties have developed have not been taken up by the AMA Physician Consortium for measure development or by the NQF for endorsement – because of limited resources, both groups list prevalence and financial impact as criteria for projects.

How Can More Quality Measures Be Developed for Specialties?

No one is saying that quality measure development and pay-for-performance in general is impossible for the specialties. What we are saying is that the mold that worked for primary care and cardiology will not work for many of the specialties. For example, one of the most outstanding quality improvement projects of the last 20 years is the clinical outcomes data registry operated by the Society of Thoracic Surgery. Through data collection, review and reports, the database has improved patient outcomes, reduced mortality and morbidity and even reduced healthcare costs by reducing complications (although under the current payment system, all of the savings go to the hospital and do not effect physician costs and expenditures). However, under the current pay-for-performance mold, the STS outcomes project is left out because it does not fit the current mandated pattern of Level I evidence = guidelines = quality measures and it cannot be easily implemented into the current mold of submitting data to payers. The point is that when it comes to physician pay-for-performance, one size does not fit all and the system needs to have enough flexibility to recognize what worked for primary care may not work for major surgery and what works for major surgery might not work for the specialties who perform mostly outpatient surgery. Policy that includes flexibility in measure development, including definitions of levels of acceptable evidence; data collection methods, including allowing clinical data over claims data; and general plan structure will allow more specialties to participate and develop programs that truly improve patient quality.

Does Pay-for-Performance Have to be Implemented Across the Board Initially?

The answer depends on how the program is structured. If the SGR and payment cuts are addressed separately and fairly across all specialties and pay-for-performance is funded with new additional money, then it would be possible for primary care and a few other specialties to begin pay-for-performance before other specialties. However, if the system is punitive and punishes physicians who do not participate with a negative physician payment update or by reducing reimbursement to fund a pay-for-performance pool, then all physicians must have the opportunity to participate. It is not fair to cut physicians if they do not even have the ability to participate.