Once upon a time, not too long ago, long waits and delays in the office were seen as something of a status symbol. A physician who had a long wait list for appointments or a long delay in the waiting room must have been an awfully good physician to have such “demand.” Even when patients began to complain, many groups were hardly eager to tackle the problem because there was a certain amount of security accompanying the status quo.

About that time, in the early 1990s, we were managing a large primary care department for Kaiser Permanente in northern California. We had roughly a quarter of a million adult primary care patients, more than 100 physicians and 400 support staff. The one thing all of us had in common was that no one was happy, with much of the discontent centered around this issue of long waits and delays, or bottlenecks, in our system. The average wait for an appointment was 55 days, and when our patients were lucky enough to get an appointment, the likelihood that they would see their own personal physician was less than 47 percent.

Not only was the system inefficient and frustrating, but it was also costly. For example, our large backlog of appointments nearly convinced us we needed to hire additional physicians and staff that we didn’t actually need. Long wait lists created a high rate of missed appointments, meaning lost income and lost opportunity. Staff resources that should have been channeled toward patient care were being wasted on triage, phone calls and managing our large backlog of appointments. In addition, the longer care was delayed, the greater the threat to quality.

Like many offices, we tweaked the system ad nauseam. We developed new appointment types; we centralized the phones; we decentralized the phones; we tweaked the reception area; we...
n conducted “service recovery” programs; and we got out our whips and beat the physicians and staff a little harder. Of course, none of that worked. There were moments of relief, but we didn’t get the kind of improvement we knew we needed to be successful. What we finally realized was that we couldn’t tweak the system any longer. It had to be rebuilt.

Our rebuilding involved creating an access system focused on the key health care product: the doctor-patient relationship. To succeed, the access system would have to include two crucial features:

1. Continuity, meaning the system’s ability to match patients with their own personal physician. Good things happen when physicians see their own patients. For example, an unpublished quarterly report by Kaiser consistently shows that patients in every demographic group are always most satisfied when they receive an appointment with their own primary care physician, rather than being bumped to another provider.

2. Capacity, meaning room on the daily schedule. In other words, when the physician comes to the office every day, is his or her schedule already completely booked, or is there room on the schedule to meet the requests for care that come in from the physician’s population of patients on that day?

The access model we created is often called “open access,” “advanced access” or “same-day scheduling.” It has one very simple yet challenging rule: Do today’s work today. Doing so enables patients to see their own personal physician on the day they call for any problem, whether urgent, routine or preventive. In less than one year, it reduced our 55-day wait to just one day; it increased dramatically the odds of patients seeing their own personal physician, and it improved physician, patient and staff satisfaction. We are even gathering evidence that it improves clinical outcomes as well. (See “Success stories,” page 48.)

When we first embarked on advanced access, we believed the model would work only in managed care environments. But as we have worked with organizations throughout the United States, Canada and Europe testing and refining these principles, we have discovered that advanced access works equally as well, if not better, in fee-for-service environments. These principles have incredibly broad applications and hold tremendous promise for physician practices of all shapes and sizes.

Three options

Every system is perfectly designed to get the results it produces. This is a key premise of the quality improvement movement, which stems largely from the work of W. Edwards Deming. What it means is that if we have a six-week delay for patients needing appointments, it’s because that’s the way we have built the system. In turn, if we want to change the result, we must change the system — explode the paradigm and look at things differently.

There are essentially three types of systems along the access continuum:

In a traditional model, each morning, the doctor goes to the office and the schedule is full. Not just full, it’s saturated, and each of the patients on the schedule made an appointment two weeks ago, a month ago, two months ago, etc. Routine appointments fill the schedule, and urgent cases are squeezed in by double-booking, skipping lunch, working late or running behind. In other words, the way practices gain capacity in these systems is to pile visits on top of an already-full schedule; they gain capacity on their backs. In a vain attempt to control demand, practices create a vast array of restrictive and complex appointment types (e.g., a female physical, a male physical, a return diabetes, a follow-up). These systems typically have high no-show rates. In addition, because schedules are full, these systems lead to an abundance of urgent care clinics, which are costly and disrupt the doctor-patient relationship. The motto for these systems is, “Do last month’s work today.”

About six years ago, researchers began looking at this problem scientifically and discovered that demand was actually fairly predictable. A number of investigators, including Marvin Smoller, MD, arrived at the conclusion that, at the level of 10,000 patients, the demand for urgent visits will be 55 on Monday, 50 on Tuesday and 45 on Wednesday through Friday — at least in their particular situation. Armed with this information, many practices moved into a carve-out model, or
first-generation open access. Every Monday, then, the practice with 10,000 patients would “carve out” (or hold) roughly 14 urgent care slots for each of its four doctors. The rest of the slots would be booked in advance, just like the traditional model. The motto for these systems is, “Do some of today’s work today.”

While the carve-out model is an improvement over the saturated schedule, it has several problems:

• First, the carve-out model actually has very little capacity because appointments are either booked (e.g., scheduled two weeks ago) or held for same-day urgent needs. Patients calling today with nonurgent needs for care continue to be pushed into the future, so work is being delayed.

• Second, these systems tend to create a third appointment type for patients who cannot be seen today but cannot wait until the end of the queue. This makes the system more complicated and eventually extends the practice’s waiting time.

• Third, there’s always tension between the routines and the urgents. How many urgents should a practice carve out for its specific situation? Is it 45 or 55 on a Monday? Precision is important because if the practice carves out slots for urgent care and does not use them, it has wasted appointment slots.

• Fourth, support staff working in these systems have a tendency to tell patients to call back on the day they want to be seen. This sabotages any attempts at predicting demand, impacts the phone system and creates a future scheduling problem.

• Finally, there’s what we call the underground economy: the pressure to steal from future “held” appointment slots to accommodate patients who don’t fit into the currently complex scheduling systems. When practices do that, the wait list begins to grow and they eventually return to a saturated model.

The third type of access system, advanced access, is far simpler, but it requires a paradigm shift. To succeed, physicians must suspend what they have thought forever. In health care, it is genetically encoded that “if you are really sick, we will see you today; if you are not really sick, you can wait.” Advanced access eliminates the distinction between urgent and routine and requires that practices “do all of today’s work today.” That’s the motto.

On Monday, then, when a physician begins the day, approximately 65 to 75 percent of his or her appointment slots are completely open. The booked slots are filled with patients who couldn’t make it in on Friday and chose Monday instead or patients whom the physician deliberately scheduled today for follow-up, referred to as “good backlog.” When patients call the physician’s office, his or her staff simply offer an appointment for today, regardless of the reason for the visit. The first question the scheduler asks is not “what is your medical
problem?” but “who is your primary care physician?” On Monday, the practice does all of Monday’s work, which means Tuesday is open, and so on.

As it turns out, demand is not insatiable, as many practices believe. Even in systems that are out of control, the demand coming from the population (and being stored in the appointment backlog) is often equal to the number of patients actually being seen daily. When physicians realize that supply and demand are in equilibrium, they see that it is indeed possible to do today’s work today.

When practices enact such a system, several magical things occur:

• First, the wait time for a routine appointment is today. No one can beat that.
• Second, practices no longer have to hold appointments in anticipation of same-day needs, so they’ve maximized their schedules and gained capacity (or appointment availability) they didn’t have before.

Success stories

In the seven years since the authors first began experimenting with advanced access, they have seen the principles result in significant improvements in office-based care.

**Kaiser Permanente, Roseville, Northern California,** the clinic in which the authors originated their work:

• Reduced the wait time for routine appointments from 55 days to one day in less than one year.
• Increased patient satisfaction scores to among the highest in the organization.
• Increased patients’ likelihood of matching with their personal physician from 47 percent to 80 percent.
• Decreased the number of visits per patient per year to 10 percent below the baseline for the previous year.

**HealthPartners Medical Group and Clinics, Bloomington, Minn.:**

• Reduced the wait time for routine appointments from 26 days to one day in just five months.
• Improved patient satisfaction; percent of patients who said they “strongly agree” that they are able to schedule an appointment within a reasonable time went from 32 percent to 58 percent; percent of patients highly satisfied went from 40 percent to 60 percent.
• Increased the percentage of patients who matched with their own physician, from 47 percent to 62 percent.
• Have sustained all of the above gains, even while being short-staffed by two physicians.

**The Mayo Clinic’s Primary Care Pediatric/Adolescent Medicine team:**

• Reduced the wait time for routine appointments from 45 days to within two days.
• Reduced the number of daily visits on average.
• Despite very different practice styles, all of the physicians were equally successful in working down their appointment backlog.

**The Alaska Native Medical Center:**

• Reduced the wait time for routine appointments in family medicine and pediatrics from over 30 days to one day.
• Increased the percentage of patients who matched with their own physician, from 28 percent to 75 percent.

**Fairview Red Wing Clinic, Red Wing, Minn.:**

• Reduced the wait time for routine appointments from nine days to one day in just eight months.
• Reduced patients’ cycle time through the office from 75 minutes to 40 minutes while increasing face-to-face time with physician.
• Increased patient, staff and provider satisfaction.
the initiative and begin working out the kinks that will be unique to your group.

The process of moving along the access continuum involves five high-leverage changes:

1. **Commit to how the practice is going to gain capacity.** In a saturated model, practices gain capacity by piling work on top of the work that's already there. In a carve-out model, practices gain capacity by holding space in anticipation of same-day needs. In an advanced-access model, practices gain capacity by doing all of today's work today, which creates maximum capacity for tomorrow. (See “True capacity,” page 47.)

2. **Reduce the backlog of appointments.** Reducing backlog (or “appointment debt”) is without question hard work and will probably involve seeing more patients each day for about six to eight weeks. It may help to mark a target date on the calendar and agree as a group that visits will not be pre-scheduled beyond that date.

   In addition, physicians can help reduce backlog by maximizing their time with each patient, looking to see whether the patient has other appointments in the near future and asking, “Can I do more with today’s visit?” For example, if Mrs. Jones is being seen for a sore throat, the physician might say, “I see you’re scheduled for a blood pressure check in two weeks. Let’s take care of that today.”

   Another strategy is to question the frequency with which physicians bring patients back to the office for follow-up. These intervals are often based on habit or culture rather than clinical importance. If physicians begin to challenge them, they will likely gain capacity in their schedules.

   Remember that there is good backlog and bad backlog. Good backlog involves two kinds of patients: 1) those who don’t want an appointment today (no more than 25 percent of patients, in our experience) and 2) those whom the physician elects to see on a specific date for follow-up, based on clinical necessity. Bad backlog, on the other hand, would involve any other patient that the practice deflects into the future.

3. **Use fewer appointment types.** Scheduling systems work best when they are stripped of their complexity and layers of rules, which lead only to error and confusion. Appointment types in advanced access are reduced to three: P (personal: your patient seeing you), T (team: your patient seeing someone else on your clinical team in your absence) and U (unestablished: for patients who are not linked with a particular physician). The appointment lengths are also going to be standardized at about 15 to 20 minutes, or the average length for the practice’s physicians, with doubled amounts used only as necessary, for lengthy procedures, physicals, etc. While actual appointment lengths certainly vary, the doctor-patient face-to-face time is actually quite short and, in most cases, can be handled in the standard 15 to 20 minutes.

4. **Develop contingency plans.** Because demand is not entirely predictable and a practice could conceivably face a day in which demand is beyond its means, groups need to develop contingency plans. When a practice makes the decision to do all of today’s work today, the first contingency is defining “today.” For most practices, it’s from about 8 a.m. until early evening. The standard pattern of demand is that it increases very quickly in the morning, flattens at about 10 a.m., drops over lunch, and then drops precipitously from about 2 p.m. on through the afternoon. The demand for appointments after 4 p.m. constitutes about 4 percent of total demand per day. (Note that when physicians see patients in the late evening and night, often that demand was created earlier in the day but was deflected when the practice did not have earlier open appointments.)

   Contingency plans should also address who else on the team can help the physician during times of excessive demand or when he or she must leave the office unexpectedly. For example, can a nurse handle additional parts of the visits? Is there a midlevel provider or colleague available (keeping in mind that the practice needs to protect doctor-patient continuity as much as possible)? Are exam rooms ready to go, fully stocked and laid out in the same way, so that physicians’ time with patients will be more efficient? Groups should also plan ahead for how they will handle predictable increases in demand. For example, a practice may conduct a flu shot campaign in
early fall to reduce demand during flu season.

5. **Reduce demand for unnecessary visits.** Cleaning up a practice’s scheduling system also involves getting rid of visits that are of little value. Even in a fee-for-service environment, this makes sense. As we mentioned earlier, one way to reduce future demand is to maximize today’s visit. Often, that means a richer visit, which is ultimately more satisfying to the patient and brings with it a higher CPT code. In turn, the capacity gained in a physician’s schedule by reducing unnecessary visits can be used to add ancillary service or more patients to the practice.

Of course, there are still other ways to reduce demand: Learn how to use the telephone and e-mail more effectively; conduct group visits; use improved care models to take care of patients with chronic illnesses; and create an effective telephone advice system that’s seen as a service, not a barrier. Perhaps most important, make sure the practice is doing the right thing the right way the first time, which is more likely to happen when patients are seeing their own physicians.

**Advanced access tips**

1. Move toward advanced access by working down your backlog of appointments.
2. Roll out the new system by showing, not telling, patients how it works. When we try to explain our systems, we often make them overly complicated.
3. Begin offering every patient an appointment on the day they call your office, regardless of the reason for the visit.
4. If patients do not want to be seen on the day they call, schedule an appointment of their choosing. Do not tell them to call back on the day they want to be seen.
5. Allow physicians to pre-schedule patients when it is clinically necessary (“good backlog”).
6. Reduce the complexity of your scheduling system to just three kinds of appointments (personal, team and unestablished) and one standard length of time.
7. Make sure each physician has a panel size that is manageable, based on his or her scope of practice, patient mix and time spent in the office.
8. Encourage efficiency and continuity by protecting physicians’ schedules from their colleagues’ overflow.
9. Develop plans for how your practice will handle times of extreme demand or physician absence.
10. Reduce future demand by maximizing today’s visit.

**Setting appropriate boundaries**

A group’s ability to do all of today’s work today and to have doctors accountable for their own patients is directly dependent upon panel size. If physicians have a panel of 500, they can do anything. If physicians have a panel of 5,000, they will continuously disappoint their patients. The panel size appropriate for an individual physician will depend on several factors: how often the doctor is in the office, the risk associated with caring for the specific panel of patients and the physician’s scope of practice. Each environment is a bit different, but the panel size for a full-time family physician taking care of his or her own patients in a mature system can be up to about 2,500.

For advanced access to succeed, it is also necessary to protect the doctor-patient relationship and individual doctor’s schedules. It’s common that when doctors make advanced-access improvements and begin to gain capacity in their schedules, they are almost immediately forced to absorb the overflow of their colleagues. This creates the wrong incentives. Instead, groups should develop guidelines to place reasonable boundaries around physicians’ practices. Physicians can still cover for one another during absences or times of extreme demand, but the general rule would be for each physician to care exclusively for his or her own patients.

When physicians see their own patients and do today’s work today, a sense of order and control is restored to their practices. We have found, almost universally, that when practices clean up their access systems and make them more coherent, the visit ceases to be a scarce commodity, patient anxiety goes down, demand goes down, visits become richer and physicians discover there is more capacity in their systems than they ever imagined. This creates the opportunity to spend more time on practice management, patient education, teaching or seeing more patients and growing the practice, not months into the future but **today**.

---

**By maximizing visits, practices can increase future capacity, which may make room for new patients or new services.**

**Advanced access depends on physicians having appropriate patient panel sizes.**

**Practices must also protect physicians from their colleagues’ patient overflow.**