Cracking the Code: Fixing the Crowded Emergency Department, Part 3 — Implementing the Solution

Joseph Guarisco, MD FAAEM
Chair, Operations Management Committee

Finally, the answer you have been waiting for! In this submission, part three of a series dedicated to providing a solution to emergency department crowding, I would like to suggest a work-flow and staffing model that helps solve capacity issues in the emergency department. Part one, published in the Sep/Oct 2013 issue of *Common Sense*, was intended to create the burning platform for change, a sense of urgency. We established that poor capacity management and long wait times negatively impact patient satisfaction, quality of care, financial performance, and patient risk. Most of the really important things we do in emergency medicine are time-critical, so solving problems related to crowding and throughput are fundamental to our goal of improving patient care. Part two, published in the Nov/Dec 2013 issue, explained the analytic approach that must be understood if one is to solve this problem. That approach is built on the premise that most of what happens in the emergency department is predictable. Equally important is the knowledge that predicting emergency department demand alone is not enough. One must also understand that failure to drill deeper, to tease apart demand and look at a more important element within the data set — variance — is an error in ED management. Understanding variance allows one to manage resources by creating a service delivery model that addresses the complexities of demand (complexity due to variance) with greater certainty at higher probabilities.

Assuming you have read and understand demand analytics as discussed in part two; that you recognize that staffing to average demand, as is commonly done, will lead to failure 50% of the time; and that you recognize the real solution is solving for variance at the 80th or 90th percentile rather than the median; then the next logical question to be asked is how does one do that without experiencing financial ruin? It's just math. Let me make it even simpler. If one takes demand data and deconstructs that data for variance, using any number of statistical tools including box plots, one can observe the peak demand that will occur at any given hour with 90% certainty or higher — or at any other cutoff you select — as noted in the box plot below (Figure 1).

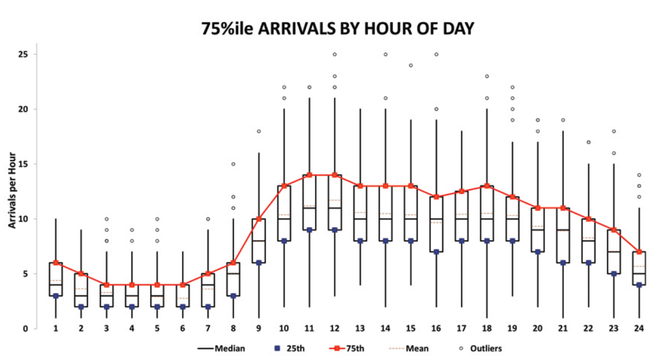

Figure 1

At the 80th percentile, just above the red line in Figure 1, there is 50% higher demand compared to the 50th percentile (black line) in Figure 2 below, if one looks across to the y-axis that represents pts/hr. The 90th percentile represents an approximate doubling of demand compared to the median, as you can see by looking at the associated pts/hr at the peak of the vertical black lines of the box plots. So, solving for the variances in demand with confidence at higher probabilities requires a significant leap in resources. The box plots allow one to solve capacity issues predictably 90% of the time instead of 50% of the time. Solving for capacity at the predicted median gets us nowhere. In fact, again, it is precisely the problem we are trying to solve.

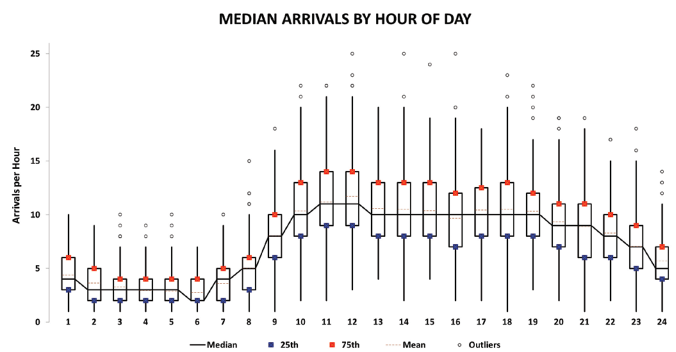

Figure 2

The obvious solution now is to allocate the necessary resources, both labor and bed capacity, to meet calculated predicted patient demand precisely. Easy solution, right? No! That solution won't work. It's too expensive.

The solution that I propose and will validate has three components: 1) a work-flow model that splits incoming patients into streams based on patient needs, 2) a staffing model that matches appropriate providers to each patient stream, and 3) optimized provider power in each patient stream using tools that are now available. The solution is to generate virtual bed capacity and increase provider productivity, while also reducing cost per visit, thus allowing management to add enough providers to service the 90th percentile probability of demand — thereby improving service performance.

Let's take each initiative one by one. Traditionally, emergency departments have one stream: patients arrive and are placed in a single queue for an emergency department bed. Patients remain in that bed until care is completed, consuming that bed for their entire visit, and then are discharged. The immediate solution seems obvious: a fast-track. A typical fast-track does achieve a second patient stream. However, as in the traditional emergency department, patients remain in that fast-track bed until care is completed, once again consuming that bed for their entire visit. The successful work-flow model must create two or three streams

Continued on next page

based on a patient's need for the most valuable resource in the department — a bed. This work-flow is shown below as a Microsoft® Visio image (Figure 3).

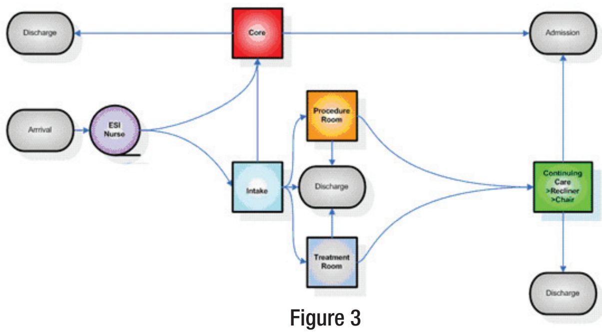

Figure 3

Patients in the second stream are managed very much in the outpatient model, in that they don't consume a bed for their entire visit, and possibly not at all. Evaluations are done, tests may be ordered and treatments administered, all in an ambulatory environment that preserves valuable beds. Our experience is that 65% of patients can be managed without using a bed. It's a simple approach but it creates valuable virtual capacity. This solution requires a functional change and sometimes an architectural change in how we stream patients through the emergency department. This is demonstrated in the work-flow design below (Figure 4).

Figure 4

Ochsner Hospital Split Flow Stream qTrack

Okay, we now have a functional work-flow that creates lots of virtual beds and saves real beds for patients who need them. Now for the second innovation, a staffing model that matches different types of providers to patients in each stream, according to skill-set and cost. Looking back again at the traditional ED, most are staffed primarily by emergency physicians, residency-trained and board-certified whenever possible. Step one, as explained above, creates capacity but doesn't guarantee staffing to demand variance at the 90th percentile. To say this another way, the traditional staffing model does not allow one to meet the service goal of short wait times 90% of the time. The assumption made in implementing step two of the solution is that patients in the newly created patient stream do not need a board-certified emergency physician. Patients in this stream are not acutely ill, though they may turn out to need comprehensive evaluation and even hospital admission. Most of these patients, however, will be ESI (Emergency Severity Index) fours and fives. Providers other than physicians are perfectly appropriate to care for this group of patients. Advanced practice clinicians (APCs, or mid-level providers) do well in this environment. We need to make a second important assumption: that 80% of the work is basic or even nonclinical, involving such tasks as electronic charting, patient navigation and escort, simple procedures, patient bonding, discharge instructions or prescriptions, etc. The diagnostic work may be fairly simple, requiring little physician oversight to ensure clinical quality and appropriateness. Most of the less critical, non-diagnostic aspects of patient care in this environment can be performed by a mid-level provider as productively and efficiently as a physician, at much lower cost. The result of this model is productively close to that of a physician, at 25-35% of the cost of a typical staffing model that meets only average patient arrival demand. As we stated earlier, however, meeting average demand is not our goal.

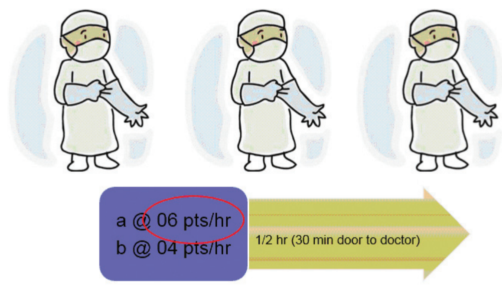

Figure 5

For demonstration purposes, using the classic queuing formula that wait time is defined roughly as 1/(a-b), where "a" is how many patients can be seen and "b" is how many patients will arrive; compared to the staffing mix above in Figure 5, the staffing mix below in Figure 6 shows that four APCs can replace one MD and achieve better service metrics at the same cost.

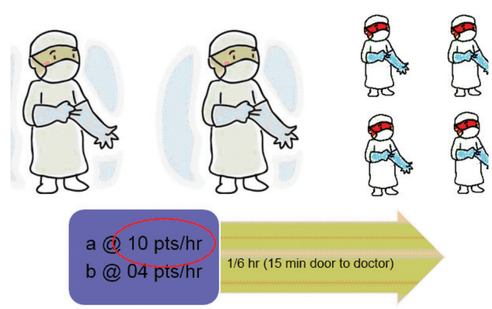

Figure 6

Again, our goal is to achieve service delivery targets such as "door to provider" at the 90th percentile of patient demand rather than the median. This staffing model allows one to increase provider "power" without adding additional cost when compared to traditional staffing. We now have both a work-flow model and a staffing model that allows us to get very close to our performance and financial goals, but we're not quite there yet. The cost advantage is demonstrated below. As you peel away physicians and add APCs, the cost per patient drops and productivity increases (Figure 7).

Continued on next page

Provider Cost Calculator							
MD	Net MD Production	MD Cost	Advanced Practice Clinician	APC Cost	Total Cost/Hr	Total Pt/Hr	Cost/Pt
2	2.0	$400.00	0	$0.00	$400.00	4	$100.00
1	0.8	$200.00	1	$50.00	$250.00	3.6	$69.44
1	0.6	$200.00	2	$100.00	$300.00	5.2	$57.69
1	0.4	$200.00	3	$150.00	$350.00	6.8	$51.47
1	0.2	$200.00	4	$200.00	$400.00	8.4	$47.62

Figure 7

Now this needs to be optimized. In other words, how many beds do you need, how many and what kinds of providers do you need, and finally, when and how do you deploy these resources? How does one make sure that the work-flow model provides for a staffing solution at the lowest possible cost per visit yet still achieves productivity and service delivery targets? How does one make sure that the mix and number of providers are yielding lower cost per visit while preserving quality of care? There are a number of expensive simulation tools that can be used for this; and there are free, tested tools in the public domain such as Banner Health Systems' Door-To-Door Toolkit, developed under a federal grant. At our facility, Ochsner in New Orleans, we use a commercially available staffing optimizer that clearly and elegantly allows one to generate a staffing model optimized for both cost and productivity, as you can see below in Figure 8.

Figure 8

Staffing Optimizer by Intrigma©

The optimizer allows one to create a provider mix (physicians, APCs, and scribes) matched to any level of demand probability. The blue line in Figure 8 is patient demand and the red line is provider power, mapped to a target of about two pts/hr in this demonstration, at a calculated cost shown in the upper left portion of the image. We have actively deployed this model and have recently added scribes to the provider mix, to further enhance physician productivity. The reduction in our costs from deploying the entire solution, when compared to the traditional emergency department management model that deploys only physicians in a single queue without optimization, is in the range of 25% — while preserving if not improving both quality and service.

In summation, in part one of this series I attempted to create the burning platform and a sense of urgency. In part two, I demonstrated an analytic model that generates a solution and avoids the pitfalls commonly made in emergency department management today, by understanding both the predictability and variability of patient demand. In this submission I propose a work-flow and a staffing model that addresses the front-end constraints that create crowding in America's emergency departments, and I propose optimization tools to solve that problem. With health care reform looming and the challenge of providing emergency care faster, better, and cheaper; this approach provides a solution for the typical ED. This solution does not diminish the need to address other throughput and output constraints, both equally important — but we will leave that for another time. I think the solution proposed in this submission is a game-changer, and goes a long way toward mitigating crowding in our emergency departments. ■