Use of Process Control to Avoid Breaching Routine Outpatient Waiting Time Targets

Dr Rod Jones (ACMA)

Statistical Advisor
Healthcare Analysis & Forecasting
Mobile: 07890 640399; www.hcaf.biz

Key Points

- Process variation sets a margin of safety for outpatient waiting times.
- The margin of safety is likely to be higher than 3 to 4 weeks for the largest clinics giving an effective 9 to 10 week average wait to guarantee a 13 weeks maximum wait.
- Smaller and less frequent clinics will have to operate at much lower average waiting times to avoid breaching the guaranteed time.
- Increasing the frequency of clinics from once a month to once a week only reduces the margin of safety by around 1 to 2 weeks.
- The government policy of a common maximum waiting time has unexpected effects based on clinic size and the sometimes unrecognised structural features of some clinics.

The UK government is fully committed to delivering its waiting time guarantees. It is fairly intuitive to see that in order to guarantee a 13 week maximum waiting time for a routine outpatient appointment must imply some margin of safety, i.e. the average routine waiting time must be lower than 13 weeks in order to avoid breaching the guarantee. How to calculate such a margin is not so intuitive.

A number of recent articles have highlighted the complexity of the factors influencing outpatient-waiting time. Monthly differences between the arrival of GP referrals (demand) and the number of first appointment slots (supply) were shown to generate a general September maximum and March minimum in number waiting and waiting time (1). Randomness in GP referrals and DNA rates has a significant effect on the average waiting time (2). The effect of randomness on the allocation of urgent appointment slots also interacts with the routine waiting time (3). Further variation is introduced by the frequency of clinics and irregular adverse events such as weather & epidemics (effecting GP referral rates) and equipment failure and consultant unavailability due to illness, holidays, meetings, etc (effecting the supply of appointments in a period).

With so many factors interacting to determine the variation in the waiting time how can a clinic be expected to avoid breaching a guaranteed maximum time? In answer to this rhetorical question this article demonstrates how techniques developed to help control

industrial processes can be used to determine the appropriate average waiting time for a clinic. This average will be below the guaranteed maximum waiting time by what is called the margin of safety.

Statistical process control was developed by the pioneering work of Shewhart (4). After observing the random variation within numerous processes in many different industries he arrived at a method which gave each process an upper and lower control limit. Anything outside these limits is considered abnormal, i.e. not due to ordinary process variation. These control limits are therefore a reflection of all the forces acting on the process and imply that a fundamental change in the process is required in order to achieve tighter control limits. In terms of outpatient waiting time all the forces acting on the process of referral, attendance and availability of routine appointments will interact to give the observed variation in waiting time.

If the upper control limit (the average waiting time plus the margin of safety) is higher than 13 weeks then the clinic will need to shift its average downward. This shift downward would need to be achieved by the one-off input of extra activity.

Having determined that a method is available to calculate a margin of safety we now need to ask the question what data should we use?

For many years almost all NHS hospitals have measured the waiting time offered to the next arriving routine referral. This forward-looking snapshot of waiting time is usually circulated to local GP's and internal management but is not reported centrally. In this instance the NHS collects information after the patient has attended their appointment, i.e. a backward looking snapshot.

In terms of management usefulness the forward-looking data is far more valuable since it indicates how long the patient will have waited when they finally attend, i.e. it allows an active rather than a passive management approach. Unfortunately few NHS managers realize the value of this information in terms of characterizing the process variation that acts to limit their ability to deliver an agreed waiting time.

The best way to illustrate is by the use of some actual clinic data. The data in Table One is a twelve month snapshot of the waiting time offered to the next arriving appointment, measured at month end, for a small (once a month – routine patients only) single consultant Orthopaedic outreach clinic held at a community hospital in Oxfordshire.

The average waiting time over the prior five-year period was 14 weeks with an upper control limit of 21 weeks. Over this period the number of GP referrals remained at an average of 110 referrals per annum although the maximum and minimum in any single year (due to Poisson randomness around the average) has been 122 and 84 respectively. The reason that referrals to this particular clinic have not changed over this period are a stable geographically defined population serviced by the same GP's over the entire period.

Shewhart control limits are calculated from the average of the absolute value of the moving range. For example, the waiting time at the end of Oct-99 and Nov-99 was 15 and 10 weeks respectively. The absolute value of the difference between these two numbers is 5, hence the moving range is 5. This calculation is then repeated for each successive pair of numbers to generate the average moving range.

Shewhart observed that the most appropriate value for the upper and lower control limits (or the margin of safety) was the average of the measurements ± 2.66 times the average of the moving range. Thus in our example the average waiting time is 11 weeks and the average moving range is 2.8 weeks, hence the upper control limits would be 18.5 weeks (11 + 7.5 weeks where 7.5 = 2.66 x 2.8). Shewhart derived the value of 2.66 by empirical observation, i.e. a value that works in practice rather than one derived from theory.

Table One: Actual clinic data and calculation of Shewhart control limits.

Month End	Waiting time (weeks)	Absolute value of moving range
Oct-99	15	
Nov-99	10	5
Dec-99	9	1
Jan-00	12	3
Feb-00	13	1
Mar-00	9	4
Apr-00	9	0
May-00	10	1
Jun-00	18	8
Jul-00	14	4
Aug-00	12	2
Sep-00	9	3
Oct-00	7	2
12 month average	**11**	**2.8**
5 year average	**14**	**2.7**

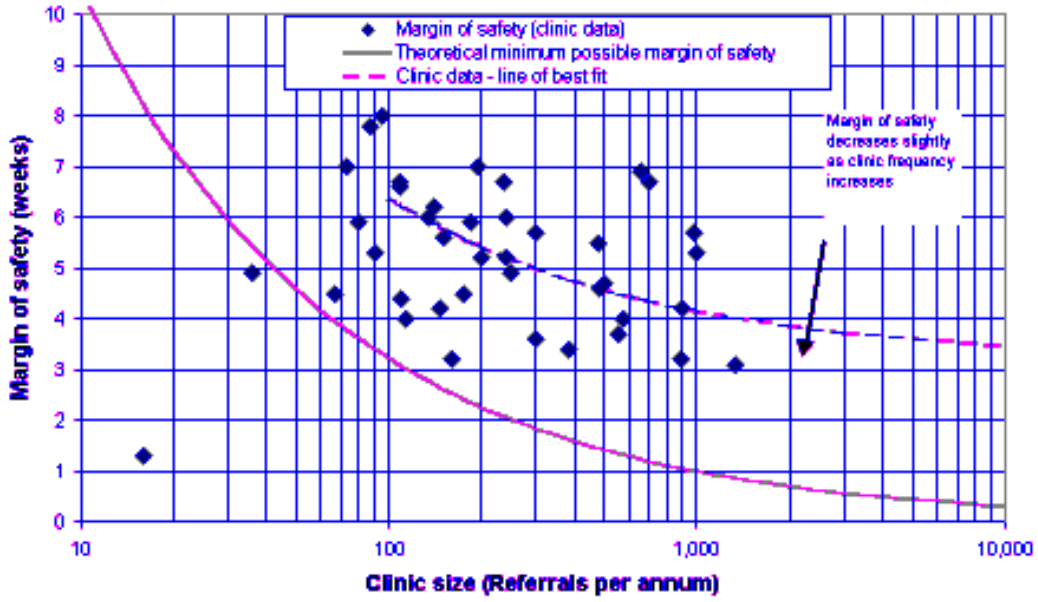
Strictly speaking, outpatient-waiting time is a ‘special’ form of process because today’s waiting time is the sum of all past events, i.e. it is dependent on the cumulative difference between referrals and number seen. Industrial processes are assumed not to show such cumulative dependence; however, in practice this effects the calculation of the average waiting time far more than it affects the value of the moving range. We can therefore calculate the value of the margin of safety using past waiting time data but need to use the most recent twelve months to calculate the current average waiting time.

In this case a twelve month average is required to deal with the known annual cycle in outpatient waiting times (1).

Returning to our central issue, namely, is it possible to deliver an agreed maximum waiting time? For the small Orthopaedic clinic the margin of safety is therefore 7.5 weeks. This implies that to avoid breaching a thirteen week waiting time this clinic would have to shift the average waiting time to 7.5 weeks lower than 13 weeks, namely to 5.5 weeks! To do this they would have to do the equivalent of 6.5 weeks work (11 – 5.5 = 6.5) as a one-off extra initiative.

Obviously clinic size will have an important effect on the margin of safety and Figure One presents the results from around 40 consultant clinics treating between 16 and 1,500 new patients per annum (5). Note that single consultant clinics rarely deal with more than 2,500 first attendances per annum (2,3).

Figure One: Clinic size and the margin of safety



Two lines have been included in the figure. The first gives the margin of safety in an ideal world where randomness in new referrals is the one and only factor influencing the outpatient waiting time, i.e. there are no complications due to seasonal patterns in GP referrals, DNA's, clinic frequency, holidays, illness, etc. This line has been calculated using computer simulation. Most real life clinics have a margin of safety higher than the ideal world simply because their margin includes all the variety of factors influencing any outpatient clinic.

The second line gives a line of best-fit through the actual clinic data. This line is approximately 3 weeks higher than the ideal case, i.e. all the other factors combine to increase the ideal case margin of safety by an additional 3 weeks.

The 'real world' clinic data shows considerable scatter around the line of best fit for four reasons. Firstly, there is some uncertainty in the calculation of the average value of the moving range simply due to Poisson randomness. This uncertainty decreases as the number of measurements increases, hence, the error associated with measurement over six months, 12 months, 24 months, 36 months and 84 months is 100%, 64%, 43%, 33% and 20% respectively. Most data came from a 24-month sample, hence the 'calculated' average moving range can be up to ± 0.75 away from the 'true' average. When 0.75 is multiplied by the factor of 2.66 recommended by Shewhart it gives a potential error of approximately ± 2 weeks for each measurement. This means that a clinic with an apparent value for the margin of safety of 3 weeks could actually have a true margin of safety somewhere up to 5 weeks.

The second reason is that smaller clinics tend to be self-regulating. For example, the two smallest clinics dealing with only 16 and 36 new referrals per annum tend to be run on an 'as needed' basis. The fact that the numbers are so small allows a totally different approach to that adopted by most other clinics (2).

Other smaller clinics dealing with around 100 new referrals per annum also have some ability for self-management in that they can slip a couple of extra patients into a clinic if waiting times look like rising. In this instance 2 extra patients represent 2% of the annual volume and hence have a significant effect on waiting time. This has the effect of reducing the margin of safety since the process for delivering waiting time is 'self-correcting' due to feedback control. The larger clinics do not have this luxury and therefore exhibit a margin of safety some 3 weeks higher than predicted for the ideal world (2).

The third reason is that some specialties experience greater monthly fluctuation in GP referrals (and hence waiting time) due to seasonal factors, e.g. varicose veins, allergy, etc (1). This automatically leads to the need for a higher margin of safety.

Lastly, variation is introduced by virtue of clinic frequency. Hence using data from the outreach clinics held in Oxfordshire where only routine appointments are seen it can be shown that an increase in the frequency of the clinic from once per month to once per week results in a 1 to 2 week reduction in the margin of safety. This relatively small reduction in the margin of safety with frequency implies that clinic size (annual volume of referrals) is the more important factor. It also implies that attempts to reduce the margin of safety by increasing clinic frequency will only be beneficial if the number of appointment slots is increased to cope with the resulting increased randomness associated with the smaller numbers per clinic (7).

At this point some would claim that the curve representing the ideal world in Figure One represents an achievable target. Simply re-design the process and the extra weeks attributed to real world clinics due to DNA's, clinic frequency, cancelled clinics, consultant illness, referral patterns, etc will be minimised.

 ☎+44(0)1276 21061

Several points need to be made regarding such an assertion. Firstly no clinic of any appreciable annual volume of referrals was able to achieve a margin of safety less than 3 weeks. This is in spite of the fact that several of the clinics are run with virtually no cancellation and most were at very low DNA rates.

The greatest contribution to the margin of safety is probably inherent in the process, namely, there are multiple process steps between the point of referral and ultimate attendance all of which contribute to a high overall process variation and hence a high necessary margin of safety. Obviously, the arrangement of cover for consultant absence due to holidays, illness, meetings and study leave is a key challenge to reducing a large part of the high process variation seen in particular clinics. However, unless the number of process steps and the multiple factors effecting these steps can be considerably reduced it is unlikely that anything better than a three-week margin of safety can be achieved.

In conclusion, for larger clinics (around 1,000 new patients per consultant firm) the theoretical minimum margin of safety is around 1 week – as defined by randomness in new referrals. This is unachievable due to the known variation in monthly patterns of GP referral (1).

Medium to large clinics can achieve an apparent minimum margin of safety of around 3 to 4 weeks. This implies that the more ‘fortunate’ NHS clinics should have an average waiting time of less than 9 to 10 weeks in order to guarantee a thirteen-week maximum wait.

Some clinics may find that structural reasons (e.g. clinics only run once a month, higher GP referrals in one portion of the year, etc) demand a much higher margin of safety and hence will have to operate at an even lower average waiting time.

Even in specialties with adequate resources the one-off extra activity required to reduce average waiting time will need the creative use of temporary rather than permanent additional resources. Small and infrequent outreach clinics will require occasional input of extra resources to counteract their high inherent variation in waiting time due to their small size and relative infrequency.

It is highly recommended that managers plot month end waiting time data on a process control chart in order to estimate when to initiate corrective action. In this instance corrective action will be required when the average waiting time for the past twelve months goes higher than 13 weeks minus the margin of safety.

References

1. Jones,R., 2000. Feeling a bit peaky, 110 (57), 28-31

2. Jones,R., 2000. A pretty little sum. Health Service Journal, 111 (5740), 28-31.

3. Jones,R. 2001. Health Service Journal, 112 (57), 28-31.

4. Wheeler,D.J & Chambers,D.S. (1992) Understanding Statistical Process Control. SPC Press Inc, Knoxville

5. Data was kindly supplied by the West Berkshire Priority Care NHS Trust and is from a large community hospital in Newbury (Berkshire) and a smaller community hospital in Henley-on-Thames (Oxfordshire).