

### 3) Interpreting a data quality report

In the previous videos we have uploaded data and created a report, and now we are going to look at a report.

When Gooroo Planner creates a report, it calculates hundreds of different things. But you don't always need to see all the detail, and usually a few highlights are going to be enough. So to do this you can attach a Report Style which picks out the things you most want to look at, and in a future video we will look at how you can customise reports styles.

When we created this report we attached a ready-made system report style which looks particularly at data quality, and that is what we're going to use here.

So we'll open this report...

And we'll use the controls to focus on a particular specialty. In this case we'll click the filter beside Specialty and select Gynaecology, and click Apply Filters, and now we are only looking at Gynaecology here in the main report view.

There are two other controls that I'll show you quickly. We can sort, either by using the long dropdown or by clicking on the header in the grey panel, and then selecting Ascending or Descending. And clicking Go.

There is one other control which we won't use here, and that is Header Grouping, and there if we unticked a header and clicked Go then it would subtotal the numbers up for us. This can be useful when you have modelled the whole trust and just want to see the headlines; however it's also worth being wary of totals because they can conceal data quality problems in the detail, so it's worth looking through the detail to check it looks sensible rather than jumping straight to the totals.

Let's look at the numbers now. This report style is all about data quality, and in particular it focuses on whether the waiting list movement data is consistent.

So if we take the first column of numbers which is new outpatients in Gynaecology, the number of additions to the waiting list (i.e. referrals accepted, in this case) is 5,099. This is the number of additions that were found in the data we uploaded, so it's measured from our own data.

The next number is an estimate of the additions, based on other data that we have uploaded. It should come to the same as the measured additions, but it almost never does. Personally I'm happy if it comes within 10%; it doesn't have to be spot on because Gooroo Planner has been designed so that minor discrepancies don't make too much difference to the overall results.

Here we can see that the estimated additions, 5,275, are only 3% different from the measured additions, so this is not going to cause any serious problems.

Then it tells us which of those two additions values it has actually used in the calculations, and here it is the measured value – you may recall from the

previous video that we asked it always to do this, by not overwriting the additions figure even if the estimated value is very different.

You are probably wondering where the estimated additions value came from, and this is explained in the next section. The logic here is that if we added 5,099 patients to the waiting list, then we should also see them cropping up in either the activity, or the other removals data, or they should contribute to any growth in the waiting list.

So if you add up the waiting list activity (4,739 new appointments), and the other removals (368 patients removed), and the growth in the waiting list (168 more patients), then that comes to 5,275 which is the estimate.

So this is where the estimated additions come from, and as long as this is within 10% or so of the measured additions, then this is acceptable. If it's out by more than 10% then it's worth trying to understand why, by studying your data, and fixing it if possible.

There are two more bits of intelligence in this report which tell us something useful about data quality.

One is the proportion of patients who are removed from the waiting list without being counted as activity. In this case 368 patients were removed, and this was 7% of the patients who were added to the waiting list, which is a perfectly reasonable sounding number as it is a fairly typical Did Not Attend rate.

The other is the visualisation of the waiting list itself, using our patient scheduling simulator. Above this service there are three icons, and the right hand one is coloured in to show us that there is patient-level waiting list data available for running the simulator.

If I click it, it launches in a separate window, and I'll click Yes to get through this screen that is warning us about a couple of appointments in the past (a common data error that you will want to correct in your patient administration system).

The symbols are explained in the Planner documentation under "Patient scheduling simulator", but briefly:

- each blob represents a patient;
- the red rectangles are clinically urgent, and the rest are routine;
- the top left quadrant is the waiting list, which shows when each patient was added to the waiting list, so the longest waiters are on the left;
- in this case the model has taken a sample (at every 5<sup>th</sup> patient) otherwise they would be too small to see properly;
- the hollow patients have not been given appointments yet;
- the solid patients have been given appointments, and if I put the mouse over one then it flashes;
- you can see another patient flashing in the bottom right, and that is the same patient, and this is when they have been booked for their appointment or procedure;
- the number of slots available each week (shown by the boxes) is based on the activity scenario we are running in this report, and the red slots are reserved for urgent patients and the blue slots are for routines.

In this case we are looking at data quality and what is most relevant is the notch in the grey bar that shows when this waiting list snapshot was taken – the census date. This notch indicates the typical rate that patients are added to the waiting list, and usually we would expect it to be roughly level with the shorter-waiting cohorts of the waiting list. In this case it is, and that's fine.

We can also scan for any other issues that would be worrying, and here I would want to look into the appointments in the past, who are shown as these patients with double letters on them; at why an urgent patient is apparently waiting 9 weeks; and why these patients are booked so many months into the future.

In this case, none of those issues are going to cause any real problems when it comes to the modelling, but it's always good to understand how errors are arising so that data quality can keep being improved.

We have just been looking at the outpatient service so far, but let's look at the elective inpatients and daycases too. Here the inpatient reconciliation is out by 35% which doesn't look good, but actually what is probably happening is that patients are being added to the daycase waiting list, so the waiting list data shows their intended management as being daycase, but in fact they may end up staying the night in hospital and then their activity is classified as inpatient.

That is why the estimated additions are apparently too large, under the inpatient service. If we consider the inpatients and daycases together then the imbalances mostly cancel each other out, and the overall difference would be only one or two percent which is not worth worrying about.

Gooroo Planner is designed to be tolerant of this kind of data issue, and differences of this kind should not cause a problem because the model doesn't rely on the additions data being accurate when it calculates demand.

That was a tour through the data quality system report style, and it's a good idea to look at this report to identify any data quality issues when you first start using Gooroo Planner.

In the next video we'll look at the most popular system report style. In a future video you will learn how to customise these report styles to your own requirements.