

# Quality and quantity

**Sarah Lambert and Katie Chadd**  
*on the importance of identifying  
and utilising quality data*

**P**revious *Research and Outcomes Forums* have demonstrated the value of using data to evaluate change. This issue, we reflect on why, in order to do this with confidence, we need to have (a) high-quality data and (b) large sets of data. We pose a clinical question and use some data from the RCSLT Online Outcomes Tool (ROOT) to demonstrate these two important factors. We will also provide some tips on what you can do to improve your own data collection.

## The power of data

Systematically collecting and evaluating routine clinical data can benefit your patients, your own clinical practice, the service you work for, and the profession as a whole. Data can be used for a number of different purposes, including lobbying for better funding at both local and national levels.

However, data is only powerful if we have enough of it and if it is of good quality. That is why all SLTs have a role to play in supporting the profession in this effort.

## Using data to explore a clinical question

**Q:** *At which stage of their education do children receiving intervention for language disorder / difficulties make most progress in therapy?*

The evidence around the optimal timing for language intervention reveals a mixed picture, and more research is required (Law, Garrett and Nye, 2003). An additional way of exploring this issue is to analyse routine clinical data, such as Therapy Outcome Measures (TOMs) (Enderby and John, 2019). The ROOT collates TOMs data from speech and language therapy services, and thus can be interrogated in this way. First, we look for all the records currently available on ROOT where the primary TOMs scale is 'developmental language disorder/difficulties (formerly child language impairment)'. This gives us 1,149 sets of data.

We need to know the age of these patients, but 291 records do not have the patient's year of birth recorded, leaving us with a much-reduced set of 858 records.

We divide up the datasets into educational age groups, using age at start of episode of care. A further four sets are removed because the age is a negative number, indicating inputting errors. There are a further 12 individuals aged zero years at the start of the episode of care. These are likely to be data entry errors, considering we are looking at intervention for language, but are not removed because we cannot be certain of this. This does, however, reduce our confidence in the dataset.

Following this data cleansing, we are left with 453 children aged 0–4 years at the start of intervention, 331 aged 5–11 years and 67 aged 12–18 (two individuals aged 20 and one aged 19 are excluded). However, not all data is completed for every domain, particularly carer wellbeing, and so when we start to analyse the data, some domains have an even smaller dataset. Again, this means our confidence in the interpretation of this data, especially in the domains where there are fewer records, is reduced.

The data was analysed to find the proportion of cases where the TOMs score went up (improvement), stayed the same (maintenance), or went down (declined), as an indication of progress made and this is displayed as a stacked bar chart in Figure 1 (right).



## REFERENCES

To see a full list of references, visit [bit.ly/BulletinReferences](https://bit.ly/BulletinReferences)



**AT A GLANCE**

# Principles for data quality



**Accuracy:** Have you entered the right patient details?



**Relevancy:** Have you established what data fields you need in order to capture the information you want to report on?



**Completeness:** Have you entered all information for the patient, and across all time points?



**Timeliness:** Have you been entering or importing data regularly, to provide current information?



**Consistency:** Have you agreed data labels with colleagues so that you can compare similar cohorts?

Adapted from Sarfin, 2021

Across all the domains, a higher proportion of children in the youngest age group had a TOMs score that increased, compared with the older two groups. This indicates that early intervention is potentially more successful than periods of intervention at later ages, although further exploration and research is needed to fully understand this. But we can also consider that the objective of intervention is likely to be different depending on age, and an improvement in the impairment domain may not be expected. For example, for an older child

at a point of transition, the aim might be to prevent the widening of the gap between the individual and their typically developing peers.

Though there is value in looking into these observations, their impact is more limited because we are analysing a relatively small number of datasets. Even though the ROOT database had records on 1,149 patients initially appearing to fit the criteria, only 74% of these could be used in analysis, with 26% being incomplete or of poor quality. In some domains this was even lower. This means our confidence in our conclusions is weakened.

We hope this example has shown the importance of data quality as well as quantity. The more accurate and complete data we have, the more confident we can be in our findings. This kind of evidence can be used to complement the findings from more traditional research, and will have implications for demonstrating the value of speech and language therapy to a wide range of stakeholders. **📧**

**SARAH LAMBERT**, RCSLT research and outcomes officer  
**KATIE CHADD**, RCSLT research manager  
 📧 root@rcslt.org

**FIGURE 1** Proportion of episodes of care showing clinically significant improvement (ie a change of +0.5 or more on the TOM), maintenance or decline between the start and end of intervention for language disorder/difficulties.

