

## NADAbase: Checking data quality

Data collection is a crucial part of service provision. It should constantly be informing the care of a client accessing your service, in real-time, at every step of their treatment journey.

Checking data quality is essential, as this ensures that your data accurately describes who is accessing your service. It also helps you to identify any gaps in your service delivery and demonstrate the effectiveness of the services provided to your clients.

### Tips for improving data quality

#### 1. *Prioritise consistency over intensity*

It can be easy to put off checking data quality. It can be tempting to leave all your data checking until one or two big sessions. But trust us – being consistent trumps intensity when it comes to data quality! Why? Two reasons.

Firstly, if you leave your data checking to one big session, you will get fatigued looking at all those numbers! It sounds basic, but checking your data in short bursts means you are more likely to spot any errors or issues. For instance, if you are checking your data quality once a month, consider checking it once a week, but for a shorter period of time.

Secondly, consistently checking your data saves you time! If you dedicate, say, 15 minutes per week to checking your data, you can identify any issues well ahead of time. But what if you leave it until just before reporting time and come across a problem? I'm feeling stressed just thinking about it!

#### 2. *Normalise checking data quality at your service*

Data is just bits of information. However, just about anyone who works with lots of data will tell you it feels like data can take on a life of its own sometimes! Especially if the data is still being collected and updated into a database regularly (like many of the Client Management Systems you are using).

Checking data quality is just a normal part of working with data. Remembering to remind yourself and other staff at your service about this can help take away some of the stress you may associate with data.

### What type of data quality issues should I look out for?

There are a few types of data quality issues to look out for. The usual data quality issues that arise are:

1. *Duplicate clients* – usually occur when an existing client is entered with a variation in the name (e.g. John Doe and Joe Doe)
2. *Duplicate episodes* – usually occur when the same episode is entered for duplicate clients
3. *Episodes opened for a long period of time* – usually occur when there has been a lull in the client's attendance to the service, and so the episode is kept open
4. *Date of birth* – usually occurs when the commencement date of the episode is placed in the date of birth section

When you see these data quality issues in your program, don't panic! Once you develop the habit of routinely checking your data quality, you'll realise that these are normal occurrences and most often, are easy fixes.

### I enter client data directly into NADAbase

To check data quality, you can use the Data Quality Check Report or the Data Quality Dashboard (click [here](#) for a short tutorial that covers both).

For the Data Quality Check Report, hover your cursor over the Reports tab in the NADAbase menu (below left). A drop-down list will provide you with a list of all the available reports – the Data Quality Check Report is sixth on the list.

The Data Quality Check Report provides a comprehensive overview of potential data entry errors and episodes that may require review (below right); it includes:

- Possible duplicated client codes
- Possible duplicated episodes
- Episodes that have remained open for more than 360 days



### Data Quality Report

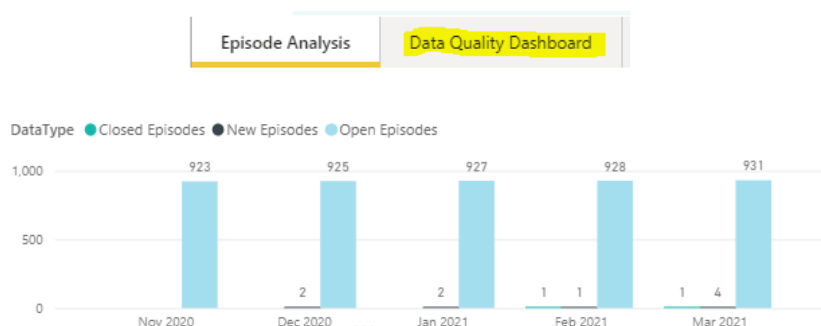
NADAbase Test Organisation  
NADAbase Test Service

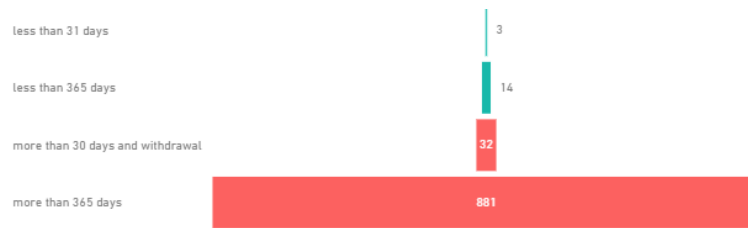
#### 3 Duplicate client codes

| Client code | Firstname | Surname | DOB        |
|-------------|-----------|---------|------------|
| Client 1    |           |         |            |
| 1           | 999       | 999     | 11/12/1980 |
| 123454321   | Jane      | Doe     | 1/01/1980  |
| 45672       | Han       | Solo    | 11/05/2000 |

The Data Quality Dashboard is accessible via NADAbase dashboard page of your program. Go to the two tabs in the bottom left-hand corner of the Episode Dashboard and click the Data Quality tab.

The Data Quality Dashboard shows useful information such as the number of closed, new and open episodes over the last five months and the number of episodes shorter or longer than 365 days.





**I don't have access to NADAbase reports. We have someone in the organisation who uploads the data for us.**

Organisations who use their Client Management System (CMS) and upload their data are called NADAbase Importers.

Usually, there are one or two people in the organisation who has access to NADAbase and uploads your program's monthly data. NADAbase has set logic and validation checks in place so it can flag any data quality issues – if it finds any issues, it won't allow the NADAbase importer to upload data.

### Import Errors

```
7 errors found:
SURVEY.txt: AgencyCode is missing on Row 2, Column 1
SURVEY.txt: EpisodeId is missing on Row 2, Column 2
SURVEY.txt: ClientCode is missing on Row 2, Column 3
SURVEY.txt: Stage is missing on Row 2, Column 4
SURVEY.txt: Submission date is missing on Row 2, Column 5
SURVEY.txt: Principal drug is missing on Row 2, Column 6
SURVEY.txt: Error on row 2:
```

There may still be a few data quality issues that get through when uploading, so it's still best practice to ask your NADAbase importer for a copy of the data quality report.