Chapter 15
Managing Concentration Data—Validation, Security, and Interpretation

Carolyn Mooney, Janice L. Abbey, and Leanne Laajoki

Introduction

The effective management of food chemical concentration data is critical when conducting a large-scale project such as a total diet study (TDS). Data validation, maintaining data security and the accurate interpretation of concentration data are all interrelated procedures required to ensure the data generated from a TDS is of high quality, accurate and representative of the food supply for the chemicals under investigation. The concentration data will underpin the subsequent estimates of dietary exposure derived for each food chemical included in the study and therefore it is important that a comprehensive data management process is followed. The approach for estimating dietary exposure will not be discussed here in any detail and is covered in depth in Chap. 17 – Dietary Exposure Assessment in a Total Diet Study.

Validation of Analytical Data

Validation of the food chemical concentration data is a process that should be conducted when the data is initially received from the analytical laboratory. This process involves scrutinizing the concentration data that is provided, often in a Microsoft Excel spreadsheet, along with all certificates of analysis to check for any errors and/or anomalies. It is important to cross-check the data recorded in the spreadsheet with all certificates of analysis to ensure consistency of reporting.

C. Mooney, B.App. Sc. (Food Science & Nutrition) • J.L. Abbey, Ph.D., B.Sc. (Hons)
L. Laajoki (✉)
Food Standards Australia New Zealand, PO Box 7186, Canberra, ACT 2610, Australia
e-mail: Leanne.laajoki@foodstandards.gov.au

G.G. Moy and R.W. Vannoort (eds.), *Total Diet Studies*, DOI 10.1007/978-1-4419-7689-5_15,

The identification of errors in the concentration data can be challenging, particularly when working with the large number of data points generated from a TDS. To assist with the accuracy of this process, the following list of errors that may be identified in analytical data can be used as a reference:

- Analytical result does not make sense:
  - An order of magnitude different to that expected
  - Unpredicted high or low values
  - Unexpected chemical for that food matrix
- Limit of Detection (LOD) or Limit of Quantification (LOQ) not low enough to provide subsequent meaningful exposure estimates for TDS purposes.
- No results reported when analysis was conducted.
- Absence of 'less than' sign as it relates to the LOQ or the Limit of Reporting (LOR), which is a trace amount between the LOQ and the LOD (see below).
- Values reported that are less than the LOQ/LOR.
- Reporting of dry weight results instead of fresh or as consumed weight results and vice versa.
- Use of incorrect units or units not reported.
- Transposition or calculation from raw instrumental data to analytical result spreadsheet.

It is recommended that two project team members working independently conduct the process outlined above in order to reduce the likelihood of errors being overlooked. Any errors or questionable results that are identified should be brought to the attention of the analytical laboratory to seek clarification. If the laboratory confirms that an error has been made in reporting, it is important that the errors are corrected by the laboratory and a revised data set forwarded to the project manager. By the laboratory making all the relevant changes, the number of individuals altering the data sheets is limited and the potential of introducing additional errors is reduced.

If the laboratory confirms that the results have been correctly reported in both the spreadsheet and the certificates of analysis, it is recommended that any available Quality Assurance (QA) information, including Quality Control (QC) data be obtained from the laboratory (See Chap. 13 – Quality Control and Assurance Issues Relating to Sampling and Analysis in a Total Diet Study). QA data provides information on the repeatability of the data on a given day using the same instrumentation (i.e. replicate analysis on the same day) and the reproducibility of the data under standard conditions (i.e. reproducibility of the data on different days by different analysts and using some altered conditions, such as reagent batches). Recovery efficiency information could also be sought from the laboratory. Recovery efficiency of food samples spiked with a known amount of the analyte, gives a good indication of the method's ability to accurately extract and detect the analyte in the food sample matrix. The concentration determined from the method is compared with the known amount that the sample was spiked with to generate a recovery efficiency. If Certified Reference Materials (CRMs) have been analyzed these results should also be checked to confirm analytical accuracy.

FOOD ANALYZED	FOOD MAPPING		
Apples e.g. Red Delicious	To all types of apples e.g. Granny Smith	To all pome fruits e.g. Nashi pear	To foods/recipes containing pome fruits e.g. apple pie
			
CONCENTRATION DATA			
Determined quantitatively by analytical methods	Quantitative concentration of analyte is extrapolated to similar foods which were not analyzed	Quantitave concentration of analyte is extrapolated to foods of a similar group which were not analyzed	A percentage of the concentration of the analyte is extrapolated to foods containing the relevant fruit (e.g. canned fruit, fruit pie and fruit juice). The percentage of the concentration used is equivalent to the percentage of the fruit in the final product. These products were not analyzed

Fig. 15.1 An illustration of validation of analytical data

After considering all QA/QC data, if there are still reservations about the results, re-analysis of the same sample by the same laboratory and/or arranging an alternate laboratory to conduct inter-laboratory check tests of the relevant samples may be appropriate. The purpose of this exercise is to reduce uncertainty around the validity of results and to allow interpretation of the data with confidence. It is important to ensure a provision is included in the contract with the analytical laboratory that requires questionable results to be re-tested and a proportion of samples to be made available for inter-laboratory check testing if considered necessary. The potential for re-analysis and inter-laboratory check testing of samples during the study should be taken into account during the sample collection stage, thus ensuring sufficient sample is collected and stored (See Chap. 9 – Food Sampling and Preparation in a Total Diet Study). Once all of the data has been checked and any errors and questionable results have been addressed, it is advised that the names of those involved in the data validation process and the date of completion are clearly documented together with any notes to the data.

Careful management of the data validation process is critical to ensure errors are not carried over into the dietary exposure assessment component of the total diet study, where their effects may be potentially amplified. For example, if *Red Delicious apples* are analyzed as part of a TDS, these values may be logically mapped to other types of apples, similar types of fruits (pome) as well as recipes containing these fruits, as illustrated in Fig. 15.1. This simple example demonstrates the importance of data validation, accurate reporting and the potential follow-on effects of an error in analyte concentration.

Security

Following completion of the data validation process, it is crucial that the spreadsheet is locked (protected) to preserve the integrity of the data. The password used to protect the spreadsheet should be created by the recipient of the original analytical data from the laboratory, generally the project manager. Protecting the spreadsheet will prevent manipulation of the raw concentration data and the potential for the introduction of errors. It is also important that the spreadsheet is appropriately named and dated, and the file is saved in a location agreed by the project team.

The TDS project team is likely to include representatives from other discipline areas within an organization, in particular the dietary exposure assessment area or equivalent. In this instance, the validated analytical data will need to be provided to the team members in this area that are responsible for completing the estimates of dietary exposure for the food chemicals investigated. It is recommended that a clearly named source spreadsheet is generated from the original validated data spreadsheet and provided electronically to the team members completing the dietary exposure assessment. By doing this, any adjustments to the spreadsheet format that are required for the purposes of calculating the estimates of dietary exposure will not affect the original data spreadsheet. This practice should preclude any issues from arising in relation to version control. It is encouraged that the procedures outlined above are clearly documented and referred to as a guide by the project team to ensure the security of the data is maintained.

Interpretation

Understanding the data and how it is obtained is essential to the interpretation process. Information such as sample composition is important, and whether the data was derived from individual or composite samples should be known. For example, if the data is derived from a composite of three primary samples and a high level of the food chemical is reported, further analysis will need to be conducted to determine whether one, two or all three primary samples are contributing to the measurement.

Understanding and subsequent interpretation of concentration data generated from the TDS is fundamental to achieving an accurate representation of the dietary exposure to chemicals from food. On occasion, the analytical data set will report results as ‘notdetected’ (ND) and ‘trace amounts’ (tr) for the analytical method. Non-detect results do not always indicate that the food chemical being analyzed is absent. In fact the chemical may be present, but its detection is limited by the sensitivity of the analytical instrument. In these cases, the food chemical would be considered as being below the LOD (Fig. 15.2). The LOD refers to the lowest concentration of a chemical that can be qualitatively detected using a specified laboratory method and/or item of laboratory equipment but cannot be accurately quantified. In contrast, trace amounts, is the term used where the food chemical has been detected by the analytical instrument (above the LOD) although the concentration cannot be quantified accurately (below the LOQ) (Fig. 15.2).

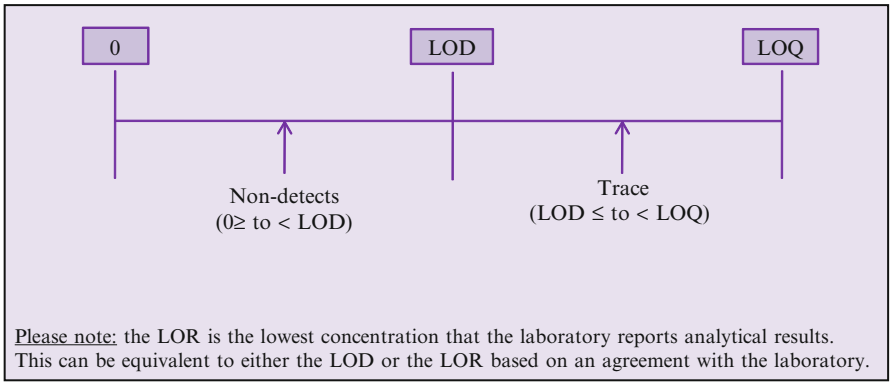

Fig. 15.2 Interpretation of non-detects and trace results in relation to the LOD and LOQ

The LOR is also a term used widely in the TDS. The LOR refers to the level of reporting which has been agreed between the project manager of the TDS and the laboratory conducting the analyses, and recorded in the contract for analytical services (see Chap. 14 – Commercial Analytical Laboratories—Tendering, Selecting, Contracting and Managing Performance). It is important to know the relationship between the LOR, LOD and LOQ, as this provides information regarding the certainty of the results. Understanding this relationship is invaluable when assigning numerical values to non-detects or trace results in order to calculate estimates of dietary exposure.

For the purposes of deriving summary statistics (e.g. minimum, mean or median and maximum concentrations) to facilitate the data interpretation process, and to allow these concentration values to inform the dietary exposure assessment, consideration needs to be given to the treatment of these results. In other words, a decision needs to be made as to what numerical concentration value to apply to non-detects and to trace results. Typically, one of the following scenarios would be applied to non-detects or trace results:

- Assigning a zero value (referred to as the lower bound)
- Assigning a value equal to half the LOQ/LOR* or LOD (referred to as the middle bound)
- Assigning a value equal to the LOQ/LOR* or LOD (referred to as upper bound)
- Assigning a range of values based on a parametric or a non-parametric method (see also Chap. 16 – Reporting and Modeling of Results Below the Limit of Detection)

*Assumes LOQ = LOR

It is important to note that the treatment of non-detects and trace results may differ depending on the type of food chemical analyzed. Factors to consider include whether the food chemical is intentionally added to food or if it is naturally present, whether both adequacy of intake and safety are being investigated (e.g. nutrients), the number of non-detect results reported and the LOQ assigned to the specific food chemical. Table 15.1 describes some of the methods used to treat non-detects in

Table 15.1 Treatment of non-detects in national total diet studies

TDS	Chemicals	Non-detects (ND)[a]
20th Australian Total Diet Study [1]	Pesticides	Reported results <LOR were included in calculation of the mean. Values <LOD were assigned 0 as pesticides are selectively applied to crops
	Metals	For values <LOR a lower bound (=0) and upper bound (ND=LOR) approach was taken and the range presented
21st Australian Total Diet Study [2]	Sorbates, sulphites, benzoates	If <LOD, ND=0 as these additives are intentionally added to food
22nd Australian Total Diet Study [3]	*Micronutrients* – iodine, selenium, molybdenum, chromium and nickel	For values <LOR, ND=1/2 LOR (middle bound) was assigned
1st French Total Diet Study [4]	Mycotoxins	For values <LOQ, ND=1/2 LOQ (middle bound) was assigned
6th New Zealand Total Diet Study [5]	*Contaminant elements* – arsenic, cadmium, lead and mercury *Nutrient elements* – Iodine, iron, selenium and sodium	As contaminant and nutrient elements are naturally occurring, ND=½ LOD was allocated

[a]Definition of ND in this context ND=<LOD or LOR, assuming LOD=LOR

national TDS reports. For example, when dealing with non-detect results, it would not be considered appropriate to apply a zero value to the food chemical if it is known to be naturally occurring in the food analyzed, as this could potentially result in an underestimation of actual concentration. For this scenario, it may be appropriate to assign a value of ½ LOQ or LOD. This is the approach that was used in the 6th New Zealand Total Diet Study for contaminant elements [5]. In relation to the treatment of trace results, for example, in the case of a nutrient for which adequacy is being assessed, applying a value equal to the LOQ could significantly overestimate the actual concentration in the foods analyzed and generate a corresponding overestimate of dietary intake. In this situation, it may be appropriate to assign a value equal to ½ LOQ. This is the approach that has been used in the 22nd Australian Total Diet Study [3].

When addressing non-detects and trace results, it is important to consider their use in exposure assessments on a case-by-case basis and ensure that any assumptions made are applied consistently and clearly documented. This will be important when preparing the final report. Once all non-detect and trace results have been considered and values assigned where necessary, summary statistics can be calculated. Reporting the median value (the statistical middle value) for the food chemicals analyzed, in addition to the mean value, may be useful where there are a large number of results below the LOD or LOR (assuming LOD=LOR) since the median is not affected by results outside the expected range. However, when there are a large number of results (n>50) and many are below the LOD or LOR, the median

cannot be calculated. In this instance, the mean is reported and the resulting assessment is conservative given the mean is higher than the median. Where there are a good number of results reported and few results reported as <LOD or LOR, it would be considered appropriate to report either the mean or median value, however the mean is more conservative. The data set is now collated and summarized and can be used to calculate estimates of dietary exposure to the food chemicals analyzed.

Summary

The management of concentration data generated from a TDS as it relates to validation, security and interpretation is vital to the quality and reliability of the study. It is recommended that procedures in relation to the format and validation of the analytical results from the laboratory should be agreed upon and stipulated in writing. Similarly, clear and detailed procedures should be in place for the project team to follow to ensure that the integrity of the data is maintained. Given that the concentration data will ultimately inform the dietary exposure assessment component of the TDS, guidance on data interpretation should be provided and consistently applied. Because conclusions regarding public health and safety will be made on the basis of analytical results, the methodical handling of such data is critical to the accuracy and representativeness of the study.

Acknowledgements The authors would like to acknowledge Mr. Peter Wallner and Ms. Jenny Trudinger for their contributions to the Australian Total Diet Study during their work at FSANZ.

References

1. Food Standards Australia New Zealand (2003) The 20th Australian Total Diet Study. Food Standards Australia New Zealand, Canberra. http://www.foodstandards.gov.au/newsroom/publications/20thaustraliantotaldietsurveyjanuary2003/index.cfm
2. Food Standards Australia New Zealand (2005) The 21st Australian Total Diet Study. Food Standards Australia New Zealand, Canberra. http://www.foodstandards.gov.au/newsroom/publications/21staustraliantotald2963.cfm
3. Food Standards Australia New Zealand (2008) The 22nd Australian Total Diet Study. Food Standards Australia New Zealand, Canberra. http://www.foodstandards.gov.au/newsroom/publications/22ndaustraliantotaldietstudy/index.cfm
4. National Institute on Agronomic Research (2004); Leblanc J-C, Guérin T, Verger P, Volatier J-L. The 1st French Total Diet Study. INRA: Institut National de la Recherche Agronomique, Paris
5. New Zealand Food Safety Authority (2005); Vannoort RW, Thomson BM. The (2003/04) 6th New Zealand Total Diet Study. http://www.nzfsa.govt.nz/science/research-projects/total-diet-survey/reports/full-final-report/nzfsa-total-diet.pdf. Accessed 17 Oct 2009