REVIEW OF THE QUALITY OF DRINKING-WATER DATA HELD IN WINZ:A DISCUSSION DOCUMENT

Prepared as part of a Ministry of Health
contract for scientific services

by

D R R Wood

May 2011

Client Report
FW11012

REVIEW OF THE QUALITY OF DRINKING-WATER DATA HELD IN WINZ: A DISCUSSION DOCUMENT

Water Programme Manager

D R R Wood
Project Leader

C Nokes
Peer Reviewer

DISCLAIMER

This report or document ("the Report") is given by the Institute of Environmental Science and Research Limited ("ESR") solely for the benefit of the Ministry of Health, Public Health Service Providers and other Third Party Beneficiaries as defined in the Contract between ESR and the Ministry of Health, and is strictly subject to the conditions laid out in that Contract.

Neither ESR nor any of its employees makes any warranty, express or implied, or assumes any legal liability or responsibility for use of the Report or its contents by any other person or organisation.

ACKNOWLEDGMENTS

The author would like thank those who were involved and helped in this work.

The author thanks all the drinking-water assessors who took the time to complete the questionnaire, the public health unit managers who gave permission for their staff to take part in the questionnaire, and the ESR staff who were interviewed as part of the data collection process. Thanks also to those individuals who provided advice and feedback on the survey design.

CONTENTS

LIST OF TABLES

LIST OF FIGURES

SUMMARY

The analysis and interpretation of public health data are needed to inform Ministry of Health policy, administer legislation and regulations, respond to questions from the Minister and Parliament, and to enable local District Health Boards and water suppliers to take public health action. However, data analysis and interpretation are only as good as the accuracy, currency and completeness of the data.

New Zealand drinking-water quality data is one such example of public health data. Water Information New Zealand (WINZ) is the hub of drinking-water quality data management and comprises 15 main modules. To be explicit about the strengths and limitations of the information held by WINZ, ESR reviewed the accuracy, currency and completeness of data within one of the modules, the supply module. Information was gathered through a workshop, interviews, an online survey, data profiling and data quality checks, to produce an assessment of the current data quality with a specific focus on the *Register of Community Water Supplies New Zealand* (the *Register*) and the water supply registration process. The focus on the *Register* and the registration process was chosen because it is the critical data set to which all water quality information is connected. This approach allowed both an in-depth look at data held on WINZ and a broader overview of the issues associated with data quality assurance. This approach could be applied to other public health datasets.

The views of drinking-water assessors (DWAs), who are key users and providers of registration information, were gathered using an online survey. This survey was sent to 31 people, had a 77% response rate and represented DWAs from across the country.

Overall, it appears that drinking-water quality data held by WINZ are reasonably accurate, though not entirely up to date, and there are a number of gaps in the information. Analysis suggests the registration-related information in WINZ corresponds with the information provided to ESR.

The main concerns around accuracy relate to population data. This is a high profile data item and is generally provided by the water supplier, with DWAs having only a minor role in population estimation. Population is far more dynamic than other registration-related information, and the size of community populations can fluctuate both seasonally and annually, usually far more rapidly than the other specifics of water supplies.

There are multiple approaches for estimating population. Comparison of the registered population information with population figures from other sources can reveal differences between the values. These differences can be interpreted as data errors, which could undermine the credibility of data held on WINZ.

There are also concerns about the definition of "specified self-supply" and how it is to be interpreted. Inconsistent, differing interpretations and/or applications of definitions such as specified self-supply, and differences in the estimation of population could lead to individual supplies not complying with Health (Drinking Water) Amendment Act 2007 and the *Drinking-Water Standards for New Zealand*. In other circumstances, these differences can make comparisons between supplies and the evaluation of national statistics in relation to populations problematic.

Some information items in WINZ are more complete than others, with the largest information gap identified being the zone GPS coordinates. This is a relatively new data item, and many supplies were registered before this item was requested by form WS01 *Application for Water Supply Registration* (it was not requested by the previous registration form). Concerns exist about how to provide a GPS coordinate (a point) for a water supply distribution zone (an area). Water suppliers and DWAs have difficulty in providing GPS coordinates in the correct format and it is generally seen as being less important than other information items.

While the information in WINZ is reasonably up to date, water suppliers are not always proactive in notifying DWAs of registration-related changes. As a result, many changes tend to be picked up during data collection for the *Annual Review of Drinking-water Quality in New Zealand*.

Though there was no evidence of any fundamental flaws with the registration process, it could be improved. The current WS01 form was noted as being user unfriendly. In addition, it does not collect all the information required to register a supply. There are other areas where the process could be modified, given better guidance or otherwise improved.

The report lays out a number of options for improving data quality, including two options that are being considered, though they have not yet been evaluated.

1 INTRODUCTION

ESR provides the Ministry of Health with national registers, information systems capabilities and access to data that are used for evidence-based assessment of the quality of drinking-water. The data are used nationally, regionally and locally to demonstrate compliance with legislation, standards and regulations. This ensures health issues relating to safe water supplies are identified and addressed effectively.

The most important piece of legislation related to drinking-water in New Zealand is the Health (Drinking Water) Amendment Act 2007 (the Act). The Act requires the Ministry to maintain a register of all drinking-water suppliers. It is the water supplier's responsibility to apply for registration and to notify the Director-General of changes to any particulars that are required. A form, WS01, *Application for Water Supply Registration* is provided by the Director-General for this purpose.

ESR supports the Ministry by managing the national drinking-water information system, Water Information New Zealand (WINZ), which contains details of over 2000 drinking-water supplies. This information is held on the *Register of Community Drinking-Water Supplies in New Zealand* (the *Register*) within WINZ. Information about drinking-water supplies on the *Register* changes daily.

An expectation of a national information system is that all users have access to it and all nationally-managed data are validated. The information held on WINZ (data analysis and interpretation) informs Ministry of Health policy, administers legislation and regulation, responds to questions from the Minister and Parliament, and it is used by local District Health Boards (DHBs) and water suppliers to take public health action, when needed. However, data analysis and interpretation are only as good as the accuracy, currency and completeness of the data. In turn, the data quality and any constraints upon its usefulness or relevance need to be conveyed to the end users when providing output.

WINZ is the hub of drinking-water quality data management, but the data flow and its quality assurance from the point of data gathering through interpretation, involves many people across many organisations. To be explicit about the quality of data held in WINZ and its constraints on analysis and interpretation, ESR has agreed with the Ministry of Health to a project that reviews and challenges the accuracy, currency and completeness of critical data held on WINZ. This is the first time that these data have been analysed using recognised approaches to assess data quality.

1.1 Project aims

The project aims to:

1. evaluate the accuracy, completeness and fitness-for-purpose of the critical drinking-water data, to identify weaknesses, and to provide direction for improvements
2. provide a data quality assessment checklist (DQAC) against which relevant ESR communications with the Ministry of Health and others can be appraised.

This report completes the first aim, providing an assessment of the accuracy, completeness and currency of one critical drinking-water data set – the particulars required for registration of a drinking-water supply under the Act. This information will be used to develop a DQAC and guide improvements in data quality assurance.

2 APPROACH

To ensure this work covered the range of factors that may influence data quality from its definition to its interpretation and use, we followed the process map illustrated in Appendix I. The reasons for using such a broad approach is that no single tool can give an overall assessment of the situation and that multiple approaches allow the results to be cross-examined. This required gathering and drawing together a number of lines of enquiry, as follows:

- ESR drinking-water team workshop
- interviews and discussions with ESR staff and a drinking-water assessor (DWA)
- literature review on data quality
- online DWA survey
- spot checks of data held on WINZ against the WS01 form
- profiling of data held on WINZ.

Figure 1 shows how these lines of evidence were brought together.

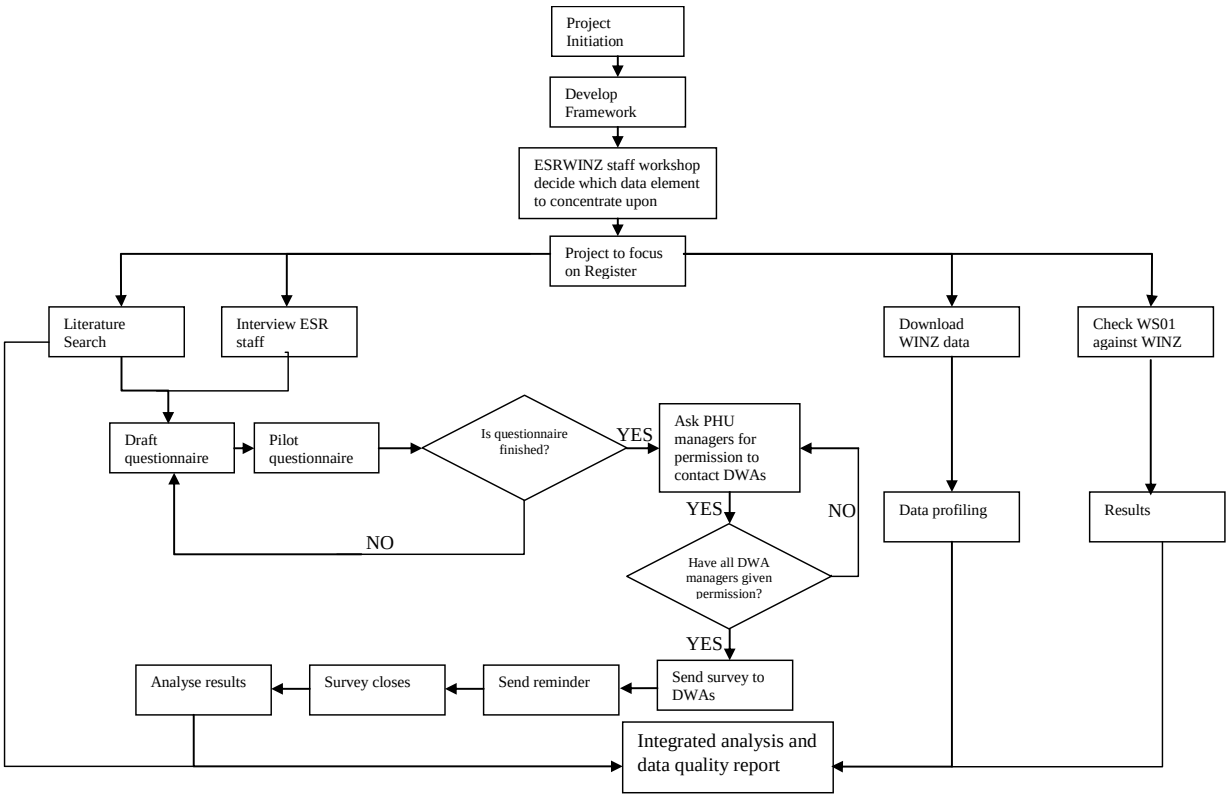

Figure 1: Process map of project indicating key steps

2.1 ESR drinking-water team workshop

The purpose of this workshop was to decide on which data set to focus. The team pooled its knowledge and experience of data gathering and used this to identify a subset of data held on WINZ where quality is critical. The criteria applied to select the data subset were:

- data were to be used widely in drinking-water management
- data collection and transfer were to be representative
- ESR to have experienced inconsistent data recording
- ESR to have experienced the impact of poor data on outputs
- provision of the data was to require a value judgment
- data were to have other definitions outside drinking-water. For example, "population" has uses and definitions beyond drinking-water.

2.2 Interviews and discussions with ESR staff and a DWA

The purpose of the interviews and discussions was to gain contextual information and understanding around drinking-water, registration processes and subjective opinions around data-quality. A series of semi-structured interviews, which allowed issues to be identified and explored, was undertaken with individuals from the ESR drinking-water team and a DWA. Prior to the interviews, drinking-water team members were asked to document the problems associated with the collection and use of registration-related information. Analysis of this document formed the interview framework.

2.3 Literature review on data quality

A brief literature review was undertaken to identify current knowledge and understanding of data quality and to apply this to the analysis of drinking-water quality data. The approach taken was intended to be informative rather than comprehensive. Documents were found through a Google search and looking through ESR files. Papers, which at first glance appeared to be relevant, were read and their key themes and issues identified.

2.4 Online DWA survey

The purpose of the online survey was to collect DWAs' opinions and views about drinking-water information, as the DWAs are key group of drinking-water information users and providers. An online survey was chosen, as it is a rapid and cost-effective method of collecting information. Permission to contact DWAs was sought from public health unit managers at each DHB. In total, 31 DWAs were invited to take part[1].

The content of the survey was informed by the interviews and the literature survey. The survey aimed to provide insight into three key dimensions of data quality, accuracy, currency and completeness. To ensure accessibility of the survey to respondents, and hence a high response rate, the format of the survey was designed to reflect their day-to-day interaction with water supply registration and WINZ.

The survey results were analysed using a mixture of descriptive univariant and multivariate statistics for quantitative data and content analysis of free text responses. Further details of the survey methodology, questions and quantitative results can be found in Appendices II–V.

[1] There are 35 people listed on the Register of Drinking Water Assessors Appointed under the Health Act 1956

2.5 Spot checks of data held on WINZ against the WS01 form

The purpose of these checks was to identify inconsistencies, if any, between data held in WINZ and the WS01 forms submitted to ESR. Eleven recently submitted WS01 forms were selected[2]. The forms were analysed to identify which items were missing or incomplete. The forms were then compared with the data held in WINZ and any discrepancies noted.

2.6 Profiling of data held on WINZ

Data profiling is the process of examining and statistically describing a database. The purpose of data profiling was to identify problems associated with registration-related information in WINZ. The WINZ database consists of a number of tables, a subset of which holds registration-related information. A copy of this WINZ subset was taken and the data examined to identify problems, such as duplicate records and/or data gaps, then a descriptive statistical analysis was undertaken.

2.7 Limitations to this approach

It should be noted that there were some limitations to the approach taken.

- This was not an audit evaluation, which would require tracking individual data items from the WINZ record back to its source. This was beyond the scope of the project.
- Data quality was viewed from the perspective of one user type, the DWAs. The issues around fitness-for-purpose (covering accuracy, completeness and currency) for other users, for example the Ministry of Health, may not be the same.
- Only a subset of WINZ data relating to the registration-related items was considered in detail, so caution must be exercised when extrapolating these findings beyond these items.

2.8 Registration process

Water supply registration data are the focus of this report. To register a supply, the water supplier and the DWA gathers the information to complete the WS01 form. Once completed, the form is signed by the person responsible for the supply and it is passed on to the DWA who signs it, and may comment on it, before passing it to ESR. For currently registered supplies, ESR then checks the information on the WS01 form ensuring it is consistent and does not conflict with other WINZ information, and that it is credible and complete. Once ESR is satisfied with the information, it is entered into the WINZ information system. Confirmation of the WINZ entry is provided to the DWA. ESR also asks whether the DWA and/or water supplier needs their copy of WINZ updated.

[2] This is a standard sample size used by inspectorates, for example the UK's Benefit Fraud Inspectorate, to indicate the presence of errors in administrative systems.

3 FINDINGS

3.1 ESR drinking-water team workshop

The workshop group decided that the review of the quality of data held on WINZ should focus on data collected for water-supply registration, because it is the critical data set to which all water quality information is connected.

3.2 Interviews and discussions with ESR staff and a DWA

The interviews and discussions provided an overview of the registration process. This included the identification of key stakeholders, and details of how information flows through, is stored and transformed, in the process of which WINZ is a part.

The interviews identified four main areas of concern, as follows:

- population data
- global positioning system (GPS) data
- the definition of a specified self-supply
- the WS01 form not requesting all the information required.

These areas of concern informed the design of the online DWA survey.

3.3 Literature review on data quality

The literature review covered ESR-Ministry correspondence (e.g. monthly reports and contracts), peer reviewed publication and grey literature.

Data quality is a performance indicator within the Ministry of Health Science Services Programme Description for Environmental Health and Communicable Diseases. It lays out the Ministry's requirements of surveillance knowledge and is included in the service description of databases systems such as WINZ.

ESR reports project progress and emerging issues to the Ministry of Health. Reviewing these reports highlighted data is fundamental to the implementation of Health (Drinking Water) Amendment Act 2007 and the operation of the drinking water assistance programme. It also identified concerns with the data and ensuring information is up to date, particularly with respect to population size. Alongside the issues with data quality, there has been an ongoing and continuous process to improve data quality. For example the design of the WINZ system has taken into account the needs of those with public health responsibilities to use the information effectively though various channels including the WINZ Advisory Group (ESR, 2004).

The peer reviewed and grey literature highlights that data quality concerns are not confined to drinking-water quality data. Assessing and managing data quality is a huge undertaking in the New Zealand health system (Kerr, 2006) and the justice sector (Ministry of Justice, 2008).

Kerr et al. (2008) noted that organisations approach data quality management in different ways. In most cases it is reactive to a data quality problem which becomes apparent when the information is used and this often leads to lack of trust in the data. A number of approaches have been proposed to assess and manage data quality such as Total Data Quality Management (TDQM), Comprehensive methodology for Data Quality (CDQ) and Ten Steps to Quality Data and Trusted Information (McGilvray, 2008; Batini et al. 2009).

Measurement is a key aspect to most approaches to improving data quality. Pipino et al. (2002) stressed the need to include both the subjective perceptions of the individuals involved with the data, and the objective measurements based on the data in question. Data quality is not just a technical information technology (IT) issue. The broader issue of data quality appears to be common to all areas that use electronically stored data.

Although the concept of "data quality" is understood, it is not well defined (Canadian Insitute for Health Information, 2005). Various definitions can be found, such as "data are fit for the purpose intended" and Kerr et al. (2008) points out that "the quality literature… describing quality in one of four general ways: as excellence, value, conformance to specifications, or meeting or exceeding consumer expectations" but any of these definitions can be difficult to apply in an operational setting.

Data quality is a multidimensional concept going well beyond just accuracy. Though there is no absolute agreement about which dimensions of data quality are important, and there is a tendency for different data quality management systems to use different dimensions (Batini et al. 2009). Some of the dimensions identified in this literature review were comparability, relevance, accessibility, consistency, believability, interpretability, understandability, accessibility, ease of operation and security, with the dimensions of accuracy and timeliness being common to five papers (Canadian Institute for Health Information, 2005; Ministry of Justice, 2008; Giovinazzo, 2009; Loshin, 2006; Ehling et al. undated). It is also clear that there is no absolute agreement about the definitions of the dimensions. For example, comprehensiveness can be thought of as the data information system containing everything it should to answer the question at hand (Coronial Services of New Zealand, Undated; London Health Services Centre, 2009), though this can be difficult to assess objectively. In addition, stakeholders have different viewpoints as data users tend to have a much wider definition of 'accuracy' than IT professionals (Wang and Strong, 1996).

The current service description for WINZ lists the requirements for registers. Registers must meet statutory requirements and be consistent with the Ministry of Health's policies, standards, guidelines and manuals. They will be accurate, complete, economical to produce, flexible, that is, they can be used for a variety of purposes, reliable, relevant, presented as simply as possible[3], timely and verifiable.

Data quality dimensions identified in the literature review that overlapped with those identified in the WINZ service description were accuracy, completeness, relevancy and timeliness or currency. Both the literature and discussions about WINZ data quality indicated that some dimensions were more significant than others. For this work the data quality dimensions focused on were accuracy, currency and completeness, which appear to be central to the most data quality concerns, including WINZ.

3.4 Online DWA survey

The online survey was sent out to 31 DWAs, and had a high response rate of 77%. Responses represented all but one DHB.

The respondents tended to be experienced, IANZ-accredited DWAs (20/23) who had worked in the drinking-water sector for longer than two years, 22/23 (96%), and 9/23 (39%) had

[3] Though without compromising other constraints.

worked in the drinking-water area for 10 years or more. They were familiar with the WS01 form, with 23/24 (96%) using it to register supplies. Of the survey respondents, 22/23 (96%) said they registered supplies and 19/23 (83%) said they modified supply registrations, a few times a year.

Some of the key responses relating to the collection and use of information are listed next.

- Overall, 20/24 (83%) of respondents thought the WS01 form collects all the information necessary to register a supply, and there were suggestions about how it could be improved.
- The forms are usually filled in by water suppliers 9/24 (38%), DWAs 9/24 (38%) or a combination of both.
- When asked to agree or disagree with various statements about whether WINZ data are accurate, up to date and comprehensive, more respondents agreed than disagreed with these statements. The most common responses were agreement about WINZ being accurate and up to date, and neutral responses for comprehensiveness (Figure 2).
- In response to questions about checking information on the WS01 form, 14/24 (58%) said all information is checked, 7/24 (29%) said some is checked and 3/24 (13%) claimed the information is not checked (Figure 3).
- Some data items collected on the WS01 forms are considered more important than other items (Table 4).
- Population is an important data item that tends to be collected by the water suppler with little input from the DWAs. Difficulties are encountered when estimating population, especially variable populations.
- Though the type of supply is important to DWAs, 11/23 (48%) said the definition of a "specified self-supply" is not clear to them, and there is confusion over the application of the definition.
- Respondents generally find out about any changes to the *Register* during data collection for the *Annual Review of Drinking-Water Quality in New Zealand* (Annual Review).
- All of the respondents agreed or strongly agreed that WINZ information is useful to them. When asked about the usefulness of WINZ data to others, the level of agreement declined in the following order: MoH, followed by drinking-water suppliers, non-health water professionals then general public and other health professionals.

Components of the results of the online survey are presented in the following sections in the context of findings from other lines of enquiry. Full details of the quantitative results of the online survey can be found in Appendix V.

3.5 Spot checks of data held on WINZ against the WS01 form

The findings from these checks are dispersed throughout the analysis section, as they have limited meaning in isolation.

3.6 Profiling of data held on WINZ

The findings from profiling the data held on WINZ are also dispersed throughout the analysis section, as they have limited meaning in isolation.

4 ANALYSIS OF ACCURACY, CURRENCY AND COMPLETENESS OF THE WATER SUPPLY REGISTRATION PROCESS DATA

Accuracy, currency and completeness of data are three key outputs of WINZ, all of which are required for its effective operation. Currently there are no agreed indicators or measures of these three dimensions. To robustly and defensibly assess the quality of the water supply registration process data in WINZ, the analysis required that multiple streams of evidence be brought together, as no single stream of evidence was sufficient. The key information sources chosen for this study were:

- DWAs, who are key stakeholders in the drinking-water management system and are central to the provision and use of these registration-related data. They have knowledge as to what is happening on the ground
- ESR staff experienced in the processing, storage and use of the data. They have a different viewpoint to the DWAs and can see how the multiple flows of information fit together
- statistical analysis of the WINZ data and WS01 forms.

Rigorous data collection approaches were selected, including online surveys of DWAs and face-to-face interviews and discussions with ESR staff. The multiple sources of information allowed results to be cross-examined, providing greater confidence in them. In addition, the integration of multiple information sources provided a more comprehensive picture of the situation.

As well as considering the outputs of data accuracy, currency and completeness, the processes that collect information converting inputs into the output were also considered.

4.1 Overview

To gain an insight into the perceived level of the overall accuracy, currency and completeness of the drinking-water quality data held in WINZ, participants in the online survey were asked in question 29 to what extent they agreed or disagreed with the following statements:

- data held in WINZ are accurate - they correctly reflect the character of a water supply
- data in WINZ are up to date
- data in WINZ are comprehensive – the data set contains all the information it should have.

The results from this question of the questionnaire are shown in Figure 2. The level of agreement, including the strongly agree and agree responses was highest for the statement "data held in WINZ is accurate...", with 17/23 (74%) of participants replying in this way, and lowest for "data in WINZ is comprehensive...", with 9/23 (39%) DWAs replying in this way. This suggests that perceived accuracy is greater than the currency, which is greater than the comprehensiveness of the data.

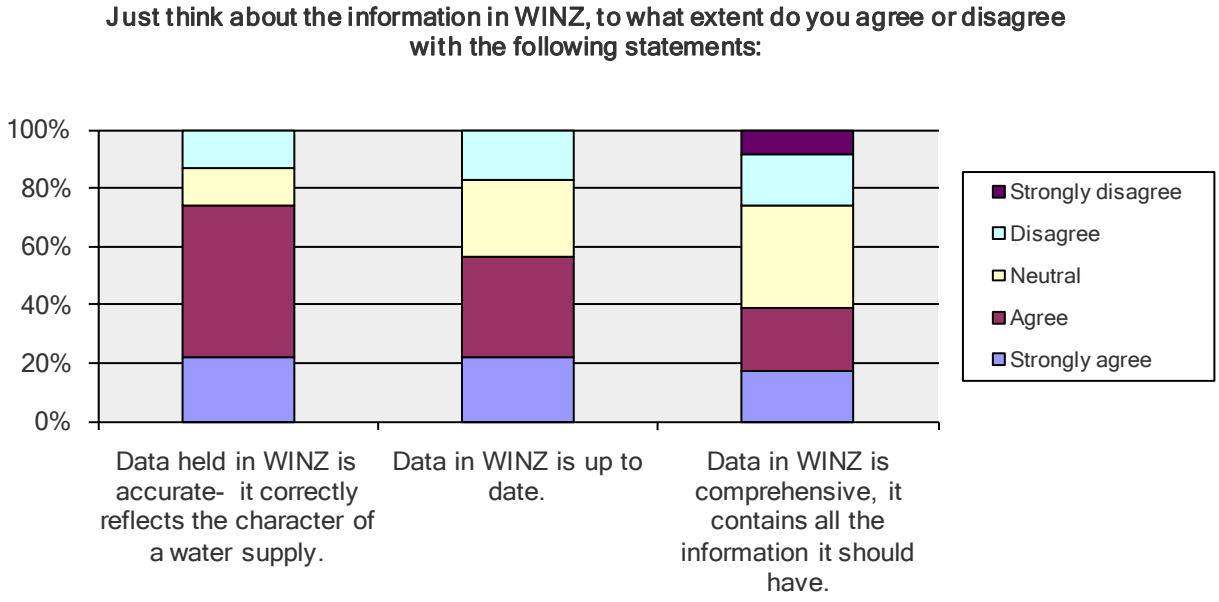

Figure 2: Indications of the accuracy, currency and comprehensiveness of information in WINZ

4.2 Accuracy

The view of accuracy in this report is determined by how well WINZ reflects the data provided to ESR. Three evidence streams were brought together, spot checks, interviews and the online survey, the latter ranking data accuracy higher than currency and comprehensiveness. This approach to data quality assessment is based on current good practice identified in the literature.

4.2.1 Comparison of WS01 with data held on WINZ

To register or change the registration of a water supply and to comply with the Act, form WS01 must be completed, and signed and submitted to ESR for entry into WINZ.

Prior to the Act coming into force, an ESR WINZ Registration form was used. The ESR WINZ Registration form collected slightly different information from the WS01 form. It should be noted that 468 of the 2300 currently registered communities (20%) have been added to WINZ during 2007 or later. Suppliers that were on the *Register* immediately before the commencement of the legislation were deemed to be registered. So it is unlikely that form WS01 has been used for these supplies, unless there has been a change to the supply.

Data quality spot checks, carried out to inform this work, compared 11 recently submitted WS01 forms against the information held in the WINZ information system. This identified two areas of difference between the WS01 forms and the WINZ records.

- Some differences in supply names between form WS01 and WINZ arose due to the naming conventions in WINZ that place restrictions on names, these include the length of name, and the need for the name to be unique and geographically identifiable (eg, “Akaroa - Bob’s café”, rather than”Bob’s café - Akaroa”). Once a name has been allocated to a supply and is recorded in WINZ, the name will remain the same unless a request to change it is made. Sometimes, the names provided by water suppliers on WS01 for new registrations or modifications to registrations are

slightly different from the name finally allocated to it in WINZ, although any changes are always made in consultation with the end users. As the key piece of information to identify a supply is the code, it was judged that these differences between WS01 and WINZ were insignificant and unlikely to cause any major problems.

- Sometimes, GPS coordinates for a zone's location had been supplied on WS01 forms, but GPS data were not apparent in the WINZ information system. In other cases, GPS coordinates were present in WINZ, but not supplied on the WS01 form being checked, which implied the information had been provided in the past. Often the GPS information on the form does not use the correct format. Differences between WS01 data and data held on WINZ appeared to be due to the GPS coordinate format rather than differences in location. It was judged, by ESR for the purposes of this report, that the GPS information, when present in WINZ, is correct, but the GPS information, overall, is incomplete. GPS information is chased up but it is generally not seen as a reason to delay the registration of a supply.

Although the checks found a couple of minor differences, it was judged that WINZ accurately reflects the registration information provided on the WS01 forms.

4.2.1.1 DWA data checks

DWAs check the information on WS01 forms. This is prior to and in addition to the checks ESR performs when it enters data into WINZ. According to the DWA survey (Figure 3), the level of data checking varies. Over half the respondents, 14/24 (58%), who answered survey question 5, said that all the data on the WS01 forms are checked for accuracy, 29% said they check some of the data, and 13% claimed that none of the data are checked.

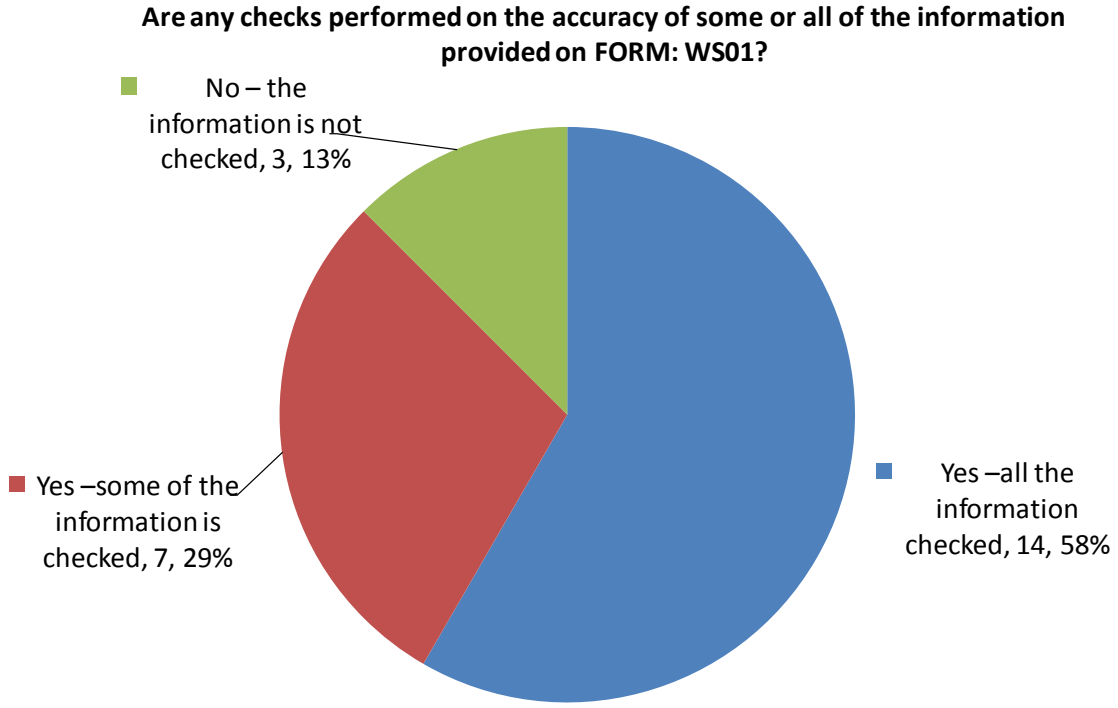

Figure 3: Accuracy checks on WS01 information

Further investigation suggested that the level of reported checking relates to who completes the forms. Survey responses suggested there is a tendency for all information to be checked when DWAs fill out forms, but only some of the information is checked when water suppliers fill out the forms. The respondents were fairly evenly split over who was the more likely to fill out WS01, with 38% suggesting the water supplier, 33% suggesting the DWA and approximately 20% claiming a mixture of the two.

When only partial checks were carried out, the responses suggested the checks depend on the situation and took the form of credibility checks. On-site visits and discussions with water suppliers are used as checks by DWAs. The main items checked included the WINZ codes, which are unique identifiers for existing registered supplies, and the GPS locations. The reason for these specific checks by DWAs according to a survey respondent was that "most applicants aren't aware what those [data items] are".

4.2.2 Issues with specific data items

Interviews with ESR staff and a DWA identified three different data areas that had issues – population data, specified self-supplies and GPS data. These issues were explored further in the online survey.

4.2.2.1 Population data

Population data held on WINZ have multiple uses. They are used to determine the level of water quality monitoring as required to comply with *Drinking-water Standards for New Zealand* (DWSNZ). Recently, population information has been used to answer drinking-water quality-related Parliamentary Questions. These data have also been used for epidemiological studies. The information for the *Register* is collected on the WS01 form and comparable data are also collected as a separate exercise for the Annual Review.

The number of people served by a water supply tends to change over time and so does the information used to estimate population data. It appears based on an understanding of demographics, the water supply process and data profiling of WINZ, that population changes more dynamically than other individual data items held on WINZ. The population figures on the WS01 forms can be estimated in a number of ways. For example, population data taken from Census figures refer to all people present in a given area on a given Census night[4]. Counts based on school rolls reflect the number of people who use the water supply during the day, and estimates based on the number of connections and residents in households reflect the number of residents in an area.

In addition to disagreements between datasets and issues about updating the information, concerns were raised during the interviews about how population figures stored in WINZ are estimated for communities. There have also been concerns that population figures are not being updated when changes occur.

The DWA survey question 13 "who usually estimates population figures?" and question 14, "how do they usually estimate population figures?", found that water suppliers usually estimate the population figures, though in some cases, usually for smaller supplies, DWAs assist water suppliers in the estimation. This response appears to be consistent with responses to a second question in the survey that asked about the DWAs' main source of advice when

[4] http://www2.stats.govt.nz/domino/external/omni/omni.nsf/wwwglsry/census+night+population+count accessed 18 January 2010

estimating supply populations. Twelve out of 23 (52%) respondents said this question "was not applicable", as they do not help to estimate the population data, 17% of respondents mentioned the legislation and/or the DWSNZ as sources of advice, colleagues were mentioned by 23% of respondents, the largest response after "Not applicable, we do not help estimate population".

All of the DWAs surveyed responded to a free text question inquiring about how population was estimated. Fifteen responses, 65% of all respondents, suggested that the estimates are based on the number of connections or houses that use a supply in conjunction with a multiplier. Census data were mentioned in 10 (44%) responses. It was also noted that the Census data tend to be used for larger supplies. For school supplies, population is estimated based on the roll.

Sixteen responses were received to question 17, "what do you see as the main difficulties, if any, in determining the population figures?". The most common response was the challenge in estimating the size of smaller populations, and particularly, estimating non-seasonal variations in population, for example, tangi on a marae. The second most commonly mentioned issue was the use of inappropriate multipliers when basing population estimates on the number of connections. Again, this is an issue for smaller supplies and, apparently at times, can grossly underestimate populations. A couple of respondents noted that as population figures determine the water quality monitoring requirements of the DWSNZ, there is pressure to underestimate populations.

When asked about which supply type/nature of supply is most difficult to define in terms of population (Figure 4), the largest response was "no opinion/no difference" by 7/23 (30%) of the people who responded to the question. This response was made up primarily of those who are not involved with estimating population data. The next two largest responses were port/airport and specified self-supply, each with 6/23 (26%) of the responses.

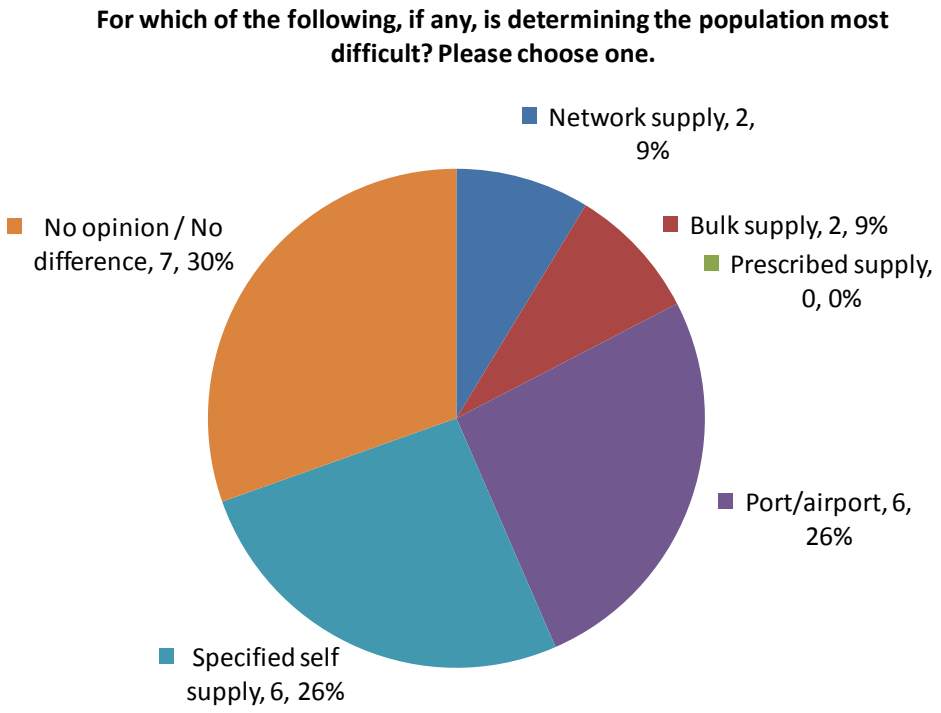

Figure 4: The supply types for which it is most difficult to estimate population

To summarise, a number of issues are experienced when estimating populations in supplies, including:

- difficulties with short-term population variations
- pressures to underestimate population figures
- no single source of information or preferred approach
- DWAs having a limited role in population estimations
- certain populations being harder to estimate than others. While these may be smaller populations, and have little impact on the total population figures, they could impact significantly on the number of registrations that are incorrect.

4.2.2.2 GPS data

Concerns have been raised about the GPS data and it is also acknowledged that there have been questions about the accuracy of GPS data. The Ministry for the Environment believes some GPS coordinates provided from WINZ to support regional councils in implementing national environmental standards are inaccurate. An investigation to assess the actual extent of any inaccuracies has yet to be undertaken, and is beyond the brief of this project.

The primary problem identified in this project appears to be around difficulties in collecting and reporting the GPS data, including reporting them in the correct format, and the data being potentially inaccurate and inadequately representing the situation on the ground. GPS data provided in the incorrect format are converted by ESR before it is entered into WINZ. These problems are considered more fully in the completeness section.

4.2.2.3 Specified self-supply

The definition of "specified self-supply" causes concern. The potential impact of this problem is an incorrect or incomplete *Register*. This difficulty is considered in the completeness section.

4.2.3 Summary of accuracy

In conclusion, there are some concerns about the accuracy of population data. Potential reasons for the inaccuracy of these data are:

- difficulties in estimating population
- multiple approaches to estimating population
- the limited role of DWAs in estimating population.

GPS data and issues around the definition of a specified self-supply are summarised in the section on data completeness.

4.3 Currency

The Service Description requires information to be "timely, ie, available when needed". The term currency describes the timeliness of data. Results from question 29 of the DWA survey suggest that currency ranked between accuracy and completeness in terms of priority/importance. As there is no agreed definition about what currency is for WINZ, evidence from the online survey and data profiling have been combined to assess the currency of the registration-related data.

As details of water-supplies change over time, information in WINZ must change in response. Data profiling describes statistically how information changes. As key users of registration-related information, DWAs can assess and provide an opinion about the currency of these data with respect to their needs. Any judgment about data currency depends on whether the information is up to date in terms of the users' specific needs and uses of the data.

Water suppliers on the *Register* must notify the Director-General of Health of a change two weeks prior to the change taking place. Evidence gathered during this exercise suggests that changes are unlikely to be notified within this time scale.

4.3.1 Is the *Register* up to date?

In terms of changes to the *Register*, survey participants were asked "How would you describe the state of the registrations in your area?" and the options were "all up to date", "mostly up to date", "somewhat up to date" and "out of date" (Figure 5). Most respondents, 18/23 (78%), said the *Register* is mostly up to date. One respondent suggested the *Register* is totally up to date, while another said it is totally out of date.

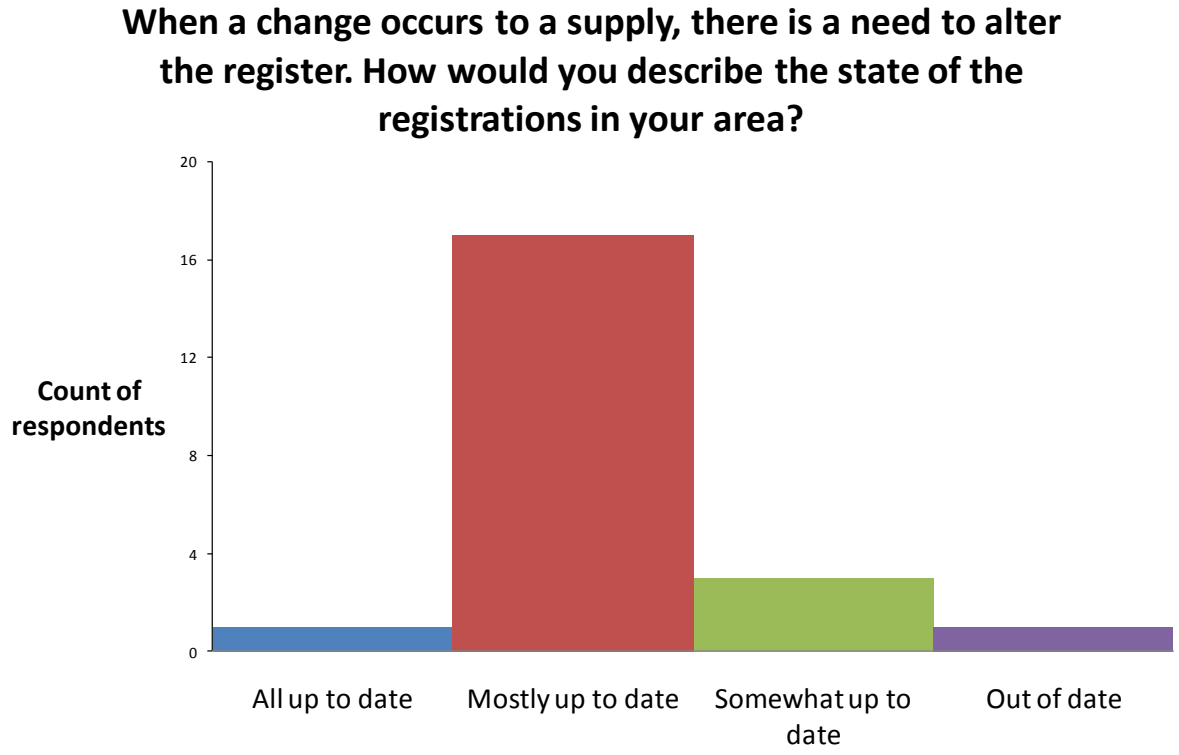

Figure 5: DWA responses to the question 'How up to date is the register in your area?'

4.3.1.1 Finding out about changes

Survey participants were asked about the way they found out about changes to the *Register*. The overwhelming response (14/23 (61%) respondents) was that changes are discovered at the time of administering the Annual Review questionnaire. Changes are also discovered through the public health grading process, public health risk management plan development and suppliers informing the DWAs.

4.3.1.2 How often do changes occur?

WINZ consists of a number of tables, four of which, the community, source, plant and zone tables, hold different aspects of registration-related information. Not all data items on the tables have to be updated at the same time to remain current.

When a record changes, a field in WINZ automatically records the date on which the last change was made. In addition, there is a notes field which can be used to provide details of the change. Data profiling was used to statistically describe the estimated average length of time since the last change for each of the four tables (Table 1). This information can be interpreted using Queuing Theory[5].

On average, the last change to each record in the community table in WINZ was 4.5 years ago. In contrast, the average last change to records in the plant table occurred 1.8 years ago.

The last recorded changes to the plant record tables were classified into four types. Of the 1992 changes to the plant record table, almost two-thirds of the changes could be classified into one of three types, population change (29%), public health grading (19%) or conversion of GPS coordinates from New Zealand Map Grid (NZMG) to New Zealand Transverse Mercator (NZTM) 2000 (17%).

The conversion of plant and source GPS coordinates to the new coordinate system was a purely technical change and did not alter any physical aspect of the supply. However, mass updating skewed the change statistics presented in Table 1. Changes to population on the other hand are in response to changes notified to ESR on a WS01 form.

Based on the community record table, information from data profiling suggests that not all WINZ records change every year. This is important as the main mechanism for DWAs to identify changes, the Annual Review, occurs more frequently, thus preventing a backlog of out-of-date records. On average we might expect 20% of the community records to have at least one change each year[5]. However, there is some evidence, based on the responses to the Annual Review, which suggests changes are not notified to ESR via the WS01 forms even when the suppliers are prompted to update or correct information.

[5] Estimate based on queuing theory (Daellenbach and McNickle, 2005) Ch 16 Waiting lines: stochastic systems.

Table 1: Average length of time for individual record changes on the WINZ supply tables

Tables	Average (mean) time for individual record changes (years)	Standard deviation for a change (years)	Proportion of active codes with a recoded modification
Community	4.5	±4.1	76%
Zone	3.1	±3.0	88%
Plant	1.8	±0.9	94%
Source	2.1	±0.9	93%

4.3.1.3 Does registering a new supply take a long time?

The survey asked "To what extent do you agree or disagree with the statement, the registration processes take a long time". Two out of 22 (9%) respondents agreed or strongly agreed that it takes a long time to register a supply, whereas 15/22 (68%) disagreed or strongly disagreed that the process takes a long time.

4.3.2 Summary of currency

The data-profiling evidence relating to currency concurs with the opinions expressed by DWAs in the survey. The main points are summarised below.

- The *Register* is mostly up to date.
- As supplies change, data items will require change.
- It is unlikely that all suppliers comply with the time specified by the Act for notifying certain types of changes.
- The Annual Review process helps to prevent the build-up of a large backlog of changes.
- The time taken for the registration process does not appear to be a significant issue in terms of data currency.

4.4 Completeness

There is no formal definition of completeness for data held on WINZ. The measure of completeness referred to in this report is the presence of data which are currently expected to be present either in WINZ or in the WS01 forms.

The results from the DWA survey suggest the completeness dimension ranks lower than either accuracy or currency in terms of priority/importance. For the purposes of the data profiling work we have defined completeness in terms of the absence of data gaps. Another way could be to think about completeness in terms of comprehensiveness. The online survey asked DWAs if data in WINZ are comprehensive, rather than complete, as DWAs do not have access to the whole of the WINZ dataset.

In this section objective measures of completeness, based on data profiling, have been used, when possible, to identify and count the completeness of data held on WINZ and the information supplied by the WS01 forms. The information was supplemented by evidence from interviews and discussions, and the online survey of DWAs.

4.4.1 Data profiling

An extract of data held in WINZ was examined during a data profiling exercise[6]. This exercise sought to describe the data, and identify duplicate records and gaps in the data. Using knowledge of the registration process it was possible to assess whether these gaps were the result of missing data, and thus the record being incomplete, or whether there was a valid reason for gaps that were not due to missing data.

Overall, the levels of completeness in WINZ, estimated by data profiling, can be ranked, in descending order as: WINZ code is more complete than the population information which is more complete than the GPS information. The above items were chosen as they are common to both WINZ and WS01. Other items are collected on the WS01 forms but not recorded on WINZ.

The data profiling exercise confirmed that each WINZ code assigned to a community, distribution zone, treatment plant or source is unique.

Every registration-related information item that needed a WINZ code had a WINZ code associated with it. Consequently, this information was judged to be 100% complete. Population data had a level of completeness that was greater than 99.9%, as shown in Table 2.

GPS data had a lower level of completeness, as shown in Table 3, with source- and plant-GPS locations having a 97% and 98% completeness level, respectively. The level of completeness of zone records is 25%. This low level of completeness reflects:

- the newness of the item[7], which was introduced with the WS01 form
- uncertainty over the meaning of the information – choice of point to represent the zone
- uncertainty about the usefulness of the information given that the zone area is already collected as a shapefile.

[6] WINZ 5.5 Supply Tables taken on 23/9/2010.
[7] Only 32% of current zone records which have changes since 1 Jan 2007 have zone GPS information.

ESR raised the issue of the zone coordinates with the Ministry of Health in the past. At the moment ESR does not chase up WS01 forms with missing zone GPS coordinates.

Table 2: Completeness of population data

WINZ table	Number of active records	Number of active records with non-zero populations	**Measure of completeness*** Proportion of records with non-zero populations
Zone	2319	2318	99.96%
Community	2300	2298	99.91%

*Data were considered complete if the code has no zero population assigned to it. The gaps found in the data have now been addressed.

Table 3: Completeness of GPS data

WINZ table	Count of active records	Count of active records with GPS data	**Measure of completeness** Proportion of records complete with GPS data
Source	2675	2598	97%
Plant	2258	2204	98%
Zone	2319	576	25%

4.4.2 **Are the WS01 forms complete?**

To help understand the level of completeness of the data in WINZ, checks were carried out on the WS01 forms. This information was analysed in conjunction with responses to a question in the DWA survey about the importance of individual data items on the WS01 form.

The DWA survey asked participants to say how important individual data items on the WS01 forms were for them. There appeared to be varying levels of importance given to items. For example, only 30% of respondents agreed that the treatment plant location description is "very important", while 87% of respondents agreed that the name of the supply is "very important".

Spot checks/data profiling on 11 WS01 forms for recently submitted new and amended registrations received by ESR, highlighted a number of information gaps. Each data item was assessed as being complete or incomplete. A data item was judged as incomplete when any aspect of it was absent; no judgment was made if the item was correct or not. The results of this test and the survey responses are presented in Table 4. Information relating to the names of supply and person responsible were complete, together with supply type and whether there were any seasonal changes in population. Other data items were incomplete to varying degrees.

A significant correlation, based on a Spearman Rank correlation, was noted between the two sets of results ($p = 0.005$). This suggests that the perceived importance of certain data items may influence their completeness. Respondents to the online survey suggest that they do not

provide information on the WS01 form, as this just duplicates information that is already in WINZ. However, for items such as GPS coordinates of zones, ESR may not have this information because it is a relatively new requirement.

Table 4: Results of the completeness checks of a sample of WS01 forms compared with the level of importance placed on individual data items by survey respondents

Data item	Proportions (%) of data items complete within 11 sample forms	Proportion (%) of respondents to survey claiming this item is "Very important"
Name of supply	100	87
Name of person responsible	100	61
Supply type	100	74
Seasonal changes in population	100	35
Name of contact person	91	61
Community name	91	61
Community population	82	61
Supplier's signature	82	NA*
Source GPS coordinates	73	32
Maximum daily volume	64	4
Distribution zone GPS coordinates	64	NA*
Treatment plant GPS coordinates	55	26
Source location description	55	39
Treatment plant location description	45	30

NA* Q25 of the survey did not include data item "Supplier's Signature" or "Distribution zone GPS coordinates"

4.4.3 GPS data

The level of completeness of GPS data is reported in Table 3. The survey asked various questions about GPS data, including question 19 "Do you provide GPS coordinates for all new registrations and modification to existing registration?", question 20 "If no, what are the barriers to providing GPS coordinates?" and question 21 "What would make it easier to provide GPS coordinates?". Overall, 83% of respondents claimed they always provide GPS data on WS01 forms, though this does not appear to be entirely consistent with the results from the analysis of the 11 WS01 forms. However, DWAs stated, in their survey responses, that they do not provide the information to ESR because they think ESR already has it. Another reason for not providing information is that it takes too long to collect – sometimes involving journeys covering hundreds of kilometres.

The survey clearly indicated there are problems with providing GPS data, and this type of information appears to be less important than other data items. Thirty-two percent and 26% of respondents suggest the source GPS coordinates and treatment plant coordinates, respectively, are very important. An added complication is that Land Information New Zealand recently changed the way the location coordinates are reported, moving from the New Zealand Map Grid (NZMG) to New Zealand Transverse Mercator 2000 (NZTM). This

may have caused some confusion, and people may be having problems reporting coordinates in the new format. The GPS information provided on the WS01 form is often in the wrong format, but the ESR converts this information and enters it into WINZ.

Suggestions made during the survey to improve the situation regarding GPS data included improved prompts on the WS01 forms, training and access to mapping programmes/databases. A few survey respondents want better access to handheld GPS equipment which is used to collect coordinate information.

Interviews and discussions with ESR staff raised the question about the definition of zone GPS coordinates. How exactly does this point location relate to the zone and the shapefile information?

The main problems associated with GPS data appear to be:

- some of the information is not provided to ESR
- the information is perceived to be of lower importance than other information provided on the WS01
- difficulties associated with providing the information in the correct format
- uncertainty as to which point (GPS coordinate) within a zone area should be given, and concern over the usefulness of providing the coordinates of a point.

4.4.4 **Specified self-supply**

Supplies applying to be registered are classified into one of five supply types, one of these being "specified self-supply". Supply classification determines whether a supply is required to be registered or not. Incorrect classification potentially leads to the *Register* being incomplete. Interviews with ESR staff suggested there is concern around the specified self-supply classification. "Specified self-supplies" are required to be registered, whereas "self-supplies" are not. There appears to be a lack of clarity as to the exact distinction between what is, and what is not, a specified self-supply.

Survey participants were asked about the source of their definition of a specified self-supply. The most widely used sources are the legislation and/or rulings from the national co-ordination team, used by 83% and 70% of respondents, respectively.

The DWA survey asked whether the definition of a specified self-supply is clear. The responses were evenly split, with 12 claiming the definition is clear, and 11 claiming it is unclear. Those DWAs who thought the definition is clear gave definitions similar to this: 'A specified self-supply is a self-supplier that provides water to buildings used for community purposes which are also owned by the self-supplier'. They also referred to a list categorising self-supplies and specified self-supplies provided by the National Drinking-water Coordination Team in consultation with the Ministry of Health, and some respondents provided extracts from that list. This group also referred to a meeting on self-supplies held in 2009 between Ministry of Health staff, national co-ordinators and ESR where the definition was discussed.

For those who do not find the definition of specified self-supplies clear, the dominant message was around the confusion and difficulty putting the definition into operation. Phrases in guidance material such as "some marae" are not considered helpful. One respondent referred to it as a "slippery pig definition".

Some comments indicated that people are not entirely happy with what the definition means, not that they do not understand the term's definition. They were concerned that applying the definition will result in some self-supplies deregistering. These supplies may currently be on the *Register* and can remain there, but they are not obliged to be registered. As mentioned earlier, this could impact on the completeness of the register. Furthermore, it raises the possibility of regional variability. It may be that some supplies are classified as specified self-supplies in one area, but not in others. This could limit data usefulness if used in regional comparison studies.

The main issues with specified self-supplies appear to be:

- confusion about definitions of a specified self-supply as opposed to a self-supply
- potential for the *Register* to be incomplete due to supplies not being registered
- removal of supplies which were on the *Register*
- potential for regional differences in the registration of supplies due to the interpretation of what is a specified self-supply.

4.4.5 Other aspects of completeness

During the interviews, ESR and DWAs noted that extra information is required on form WS01.

Some of the extra information suggested by DWAs is already collected by other means, such as the public health grading process, though obviously not all supplies currently have a grade, so is over and above what is legally required to register a supply. In addition, there was no obvious pattern between responses.

From the survey, some DWAs suggested more technical information should be collected about the supply by form WS01, and other information that local authorities deem as important should also be collected.

ESR staff suggested that information be gathered with respect to local authority areas as this information is not collected by the WS01 form. This information is currently gathered by ESR because it is required for the collation of the Annual Review data and the *Register* on www.drinkingwater.co.nz

4.4.6 Summary of completeness

The approach taken to assess completeness did not involve an audit of supplies, so it is not possible to know how well the *Register* reflects the actual situation. However, it is clear that:

- information held in WINZ reflects the information provided to ESR on WS01 forms
- some items in WINZ are more complete than others
- zone GPS information is the least complete
- there is some debate about the accuracy of the GPS information
- concerns about completeness of data may be a function of its perceived importance by those involved in data provision
- confusion around the interpretation and application of the term 'specified self-supply' may lead to the *Register* being incomplete
- some DWAs thought more information should be collected as part of supply registration.

4.5 Processes and data *quality*

Whereas accuracy, currency and completeness are the outputs of the drinking-water quality data system, it is the underlying processes that transform inputs into outputs. Knowledge of the process provides an understanding of how data quality can be changed.

The process of collecting, storing and using information and data directly involves water suppliers, DWAs and ESR. Data quality is influenced by people, processes and their interactions. Data quality can be thought of as an output of the drinking-water quality data system. In this section some process factors and how they relate to data quality are considered.

The survey asked DWAs a number of questions about the process, and the responses to these are presented in Figure 6. Overall, most respondents (18/23) thought the registration process is simple and does not take a long time. Half the respondents thought the information is easy to provide, and, overall, they were neutral 11/22 (50%) about the suggestion that the process could be improved.

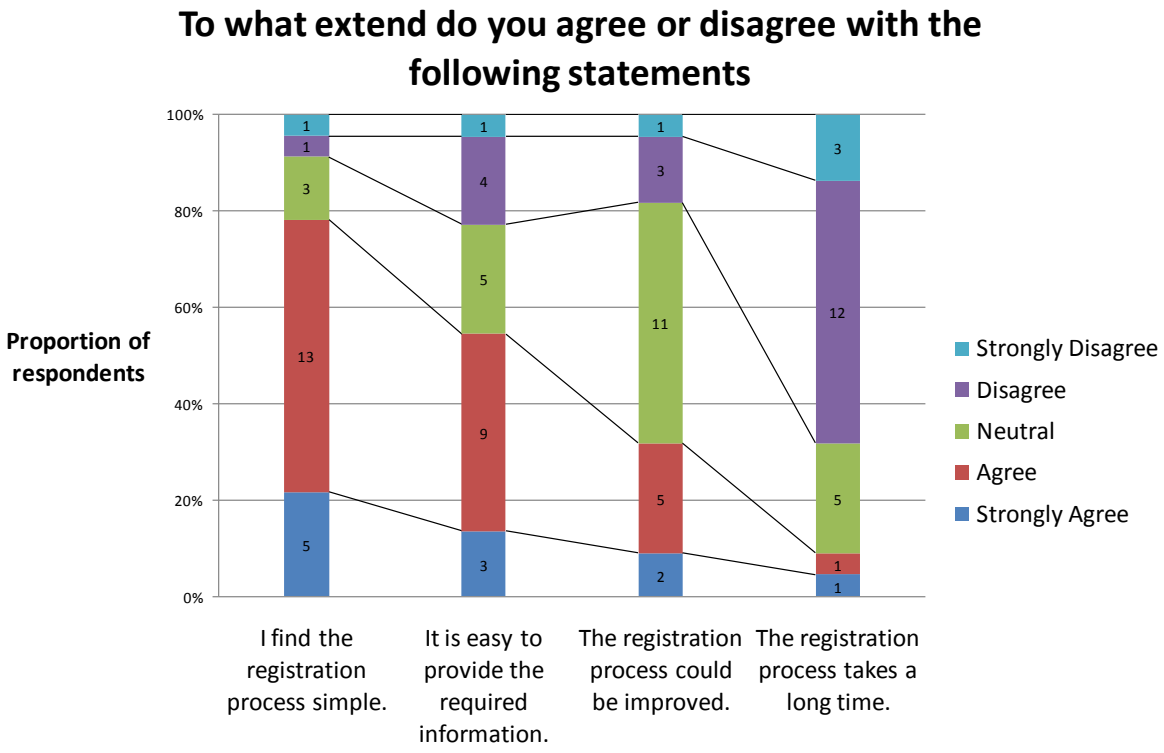

Figure 6: DWAs views about the registration process

When asked about how the process could be improved, there were nine responses. The main issues identified are:

- lack of user friendliness of the WS01 form
- limitations of WINZ
- streamlining of processes.

Problems with the WS01 form were highlighted by survey participants. A number of those who thought the process could be improved suggested that the design of the WS01 form

could be improved. One respondent suggested the current form is "for technically minded workers that do it for the DHB [not some of the water suppliers]". It was suggested that:

- the notes section, currently located on the last page of the WS01 form, is in the wrong place as these notes are often missed
- better prompts are required to help those filling in the form provide the correct GPS information
- the form needs to provide more space for the DWA's comments
- the form needs a space for the DWA's signature.

Information collected during the interview stage highlighted one other omission from the WS01 form. Registration requires the provision of the local authority area in which the supply is located. ESR collects this information in a separate step.

Respondents to the survey noted a relationship between WINZ and data quality (which is explored further below). Five respondents strongly supported the roll out of WINZ6 as they thought it would bring about process improvements including those to data quality. In addition, they say WINZ6 overcomes the data exchange problems associated with WINZ5.

Respondents to the survey highlighted that there may be opportunities to improve and or streamline the registration process. This could include exploring the use of online forms to register supplies, and identifying synergies between the functions of ESR and DWAs.

It was noted that the ESR data checking process uses high levels of expertise and competence within the organisation to perform these checks alongside the data entry function. Though this is good practice, it was also noted that the data entry and data checking processes are not fully documented as standard operating procedures. Documenting these processes would help improve quality assurance.

4.5.1 **Other factors influencing data quality**

Using factor analysis to identify latent variables[8], analysis of the survey data suggests that a number of factors influence a DWA's perception of data quality. The results of this analysis illustrate that the perceptions of data quality (associated with accuracy, currency and completeness) relate to the frequency of use of WINZ and the perceived ease of its use (Figure 7). Regular users of WINZ tend to have a more positive view about the quality of data than less regular users of WINZ. In addition, respondents who think WINZ is easy to use, have a more positive view of data quality than those who have difficulties using WINZ. Regular users of WINZ believe drinking-water information is of greater use to others (Ministry of Health officials, drinking-water suppliers and non-health-related water professions such as engineers) than less regular users of WINZ.

[8] Variables that are not directly observed but inferred through a statistical model from variables that are directly measured.

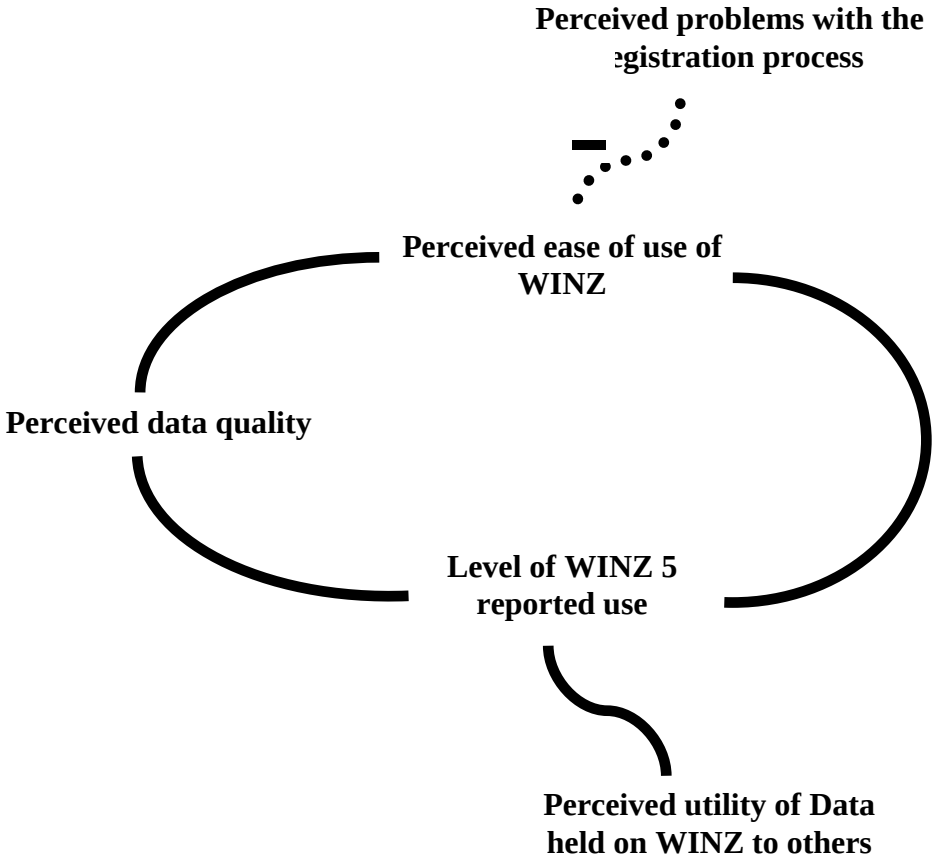

Figure 7: Influence diagram indicating the relationships between latent variables identified for survey. The negative sign by the dashed line indicates a inverse relationship, so there is a tendency for those who have higher perceived ease of use of WINZ to have fewer problems with the registration process

The relationship between WINZ use and data quality illustrates that data quality is not independent of data or other aspects of the drinking-water quality data system. This knowledge needs to be taken into account when developing data quality improvement strategies.

5 DISCUSSION

This report assesses the accuracy, completeness and fitness-for-purpose of critical drinking-water data sets of which the registration data are an example.

The data held in WINZ are an accurate representation of registration data recorded on WS01 forms. This assessment is based on comparison of the two data sources, namely, DWAs' opinions and the analysis of WINZ and WS01 forms. However, there are concerns about how well the data that are collected by the WS01 forms, and stored and processed in WINZ, reflect the actual situation. The accuracy of the population data is the primary concern.

The information held in WINZ appears to be reasonably up to date. Water suppliers should proactively notify DWAs of changes to their supply, but changes are often discovered by DWAs whilst they undertake the Annual Review. It would appear that recorded changes to supply details, on average, occur less frequently than once a year. So, assuming all changes have been recorded, the *majority* of records will be up to date. However, for some uses of the information *all* records need to be current and known to be current.

Most of the details of supplies, required for their registration, are complete, with the exception of zone GPS coordinates. At present, the impact of this missing information appears to be minimal.

There are concerns about the application of the definition of specified self-supplies. In theory, it could lead to the *Register* being incomplete, due to the possibility of misclassification of supplies. Variable interpretations of specified self-supply could also result in inconsistencies in the application of the Act to water supplies.

Overall, it appears the registration-related data held in WINZ are suitable for some, but not all, purposes for which it is currently used.

A number of factors influence the data quality and their fitness-for-purpose, including:

- data definitions and interpretation
- data collection systems
- perception and data quality
- uses of information.

5.1.1 Data definitions and interpretation

For the *Register*, in most cases, the definition of data items is straightforward and there is little room for interpretation. However, for the population data, the definition of specified self-supply and the zone GPS coordinates, there appear to be opportunities for multiple interpretations as to what the information is and how to collect it. In some cases, such as population, there may be valid reasons why different estimation methods are currently applied to data collection (Figure 8). However, this can lead to problems when the information is used, such as when national water supply statistics, relating to population, are being assembled. So, the definition of data items needs to account for the ways the information will be used.

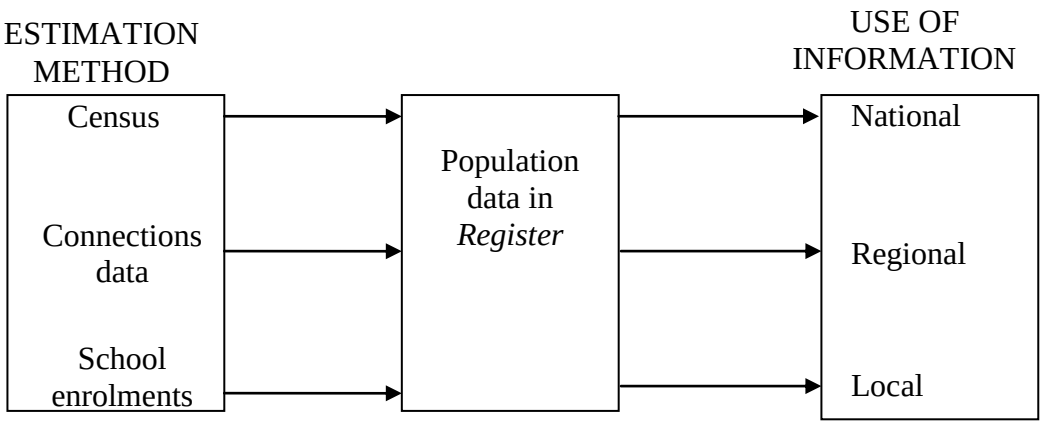

Figure 8: Multiple uses and multiple estimation methods, the challenge of accurately reporting population

5.1.2 Data collection systems

The collection, processing, storage and use of registration-related information are parts of a system, involving people across multiple organisations. A key tool used by water suppliers and DWAs is the WS01 form. It is a key component of the system and provides a focal point for discussion around supply registration and records information. However, this data collection device is described as not user friendly, and at times it may even hinder data collection.

The registration-related data collection system has a number of checks and balances within it to verify and validate information. This includes International Accreditation New Zealand accreditation of DWAs, but there is no formal quality system that crosses organisational boundaries for WINZ or the *Register*. So there is room to improve quality assurance processes, including ESR's quality assurance systems. The exact form of an improved quality system has not yet been explored or discussed, but this report represents the start of the process.

From a quality point of view, it is interesting to note that DWAs have a limited role in the collection of population data, currently one of the most problematic data items on the *Register*.

5.1.3 Perception and data quality

Difficulties with the real or perceived quality of data can influence the provision and use of information. In the absence of objective quality metrics, users of the information have to rely on their own or others' judgments as to the quality of data, its strengths and limitations. Incorrect judgements can lead to too much or too little confidence in the data.

5.1.4 Uses of information

The use of information evolves over time and so must the processes to support those uses. This may mean that information that was once fit-for-purpose may not be fit-for-purpose at some later time. At one time the *Register* was published once a year, now registration-related information is published online.

5.1.5 **Knowing whether information is up to date**

Currently the onus for keeping records up to date is with the water supplier. DWAs appear, in many cases, to find out about changes after they have happened. When discrepancies between population figures between datasets have been identified, ESR has pursued these. However, the need remains for a WS01 form that is signed by the supplier to update and modify WINZ records.

There is no way of knowing in WINZ if a record is up to date. The assumption is when no change is notified that the records are up to date. However, as time goes by there is a chance that the one or more aspects of the record could change. So should water suppliers be asked to confirm the details as correct, and should this information be held on WINZ to improve confidence in the currency of the data?

5.1.6 **Conclusion**

The challenge is to get the right information and keep the information right to support the good administration of drinking-water. There are areas where the process or data produced by the process could be improved. The information held in WINZ is fit for many of the purposes for which it is collected.

5.2 Potential improvements

Based on an analysis of suggestions made by respondents to the survey and during the interviews, and following good practice identified in the literature, options have been identified to improve data quality (Table 5). These options have not been evaluated for their feasibility, efficacy, efficiency or cost.

It should be noted, steps are already being taken to improve data quality. In response to discrepancies between the registration and Annual Review data sets, water suppliers are being encouraged to update the *Register* more regularly. In addition two options, options 1 and 2 are currently being investigated.

Option 1 – Improving population data

Discussions have already taken place with the Ministry of Health around how population data are to be collected and reported in the future. The preliminary suggestion is to base zone and community population estimates on Census mesh block data and combine these with shapefiles (shapefiles are polygons rather than a point). The details have to be worked out including how to address non-resident and variable populations. In addition, this project has highlighted potential concerns with the quality of GPS information on which the shapefiles, required for the new population estimations, would be based. If the location information is inaccurate, then any new method based on this information result in a set of data quality issues.

Option 2 – Improving quality assurance processes

ESR has noted that the quality assurance could be improved by more complete documentation of the data quality checks. These will be implemented.

5.3 Points for discussion

Data quality improvement and management is not a one of exercise, it is a continuous process. The uses of WINZ data are evolving and changing over time. Though there are already a number of procedures and activities that actively contribute to data quality assurance, there is room for improvement.

There is a question about the most appropriate way forward. Though potential options have been identified they need to be discussed, assessed and prioritised to ensure the choice of options efficiently and effectively supports the Ministry of Health's goals for water management.

Table 5: Options for data quality improvement

Area	Issue	Potential options
Population	Multiple approaches to estimating population/possible inconsistencies in approaches	• See Option 1 above
	Seasonal/variable population	• See Option 1 above
Specified self-supply	Difficulty in operationalising definition of specified self-supply	• Estimate the scale of the problem, how many supplies are potentially problematic? • Provide improved clarification of what is and what is not a specified self-supply
GPS data	Absence of GPS data	• Increase importance of GPS data, for example using it in population estimation. • Provide training to DWA in collecting and reporting GPS data after discussions and clarification with the Ministry of Health • Increase availability of up-to-date GPS and mapping resources
	GPS data in wrong format	• Provide training to DWA in collecting and reporting GPS data • Increase availability of up-to-date GPS and mapping resources
WSO1 form	WS01 form not user friendly	• Redesign WS01 form to make it more user-friendly • Develop additional guidance to assist the completion of WS01 form, possibly as a booklet to accompany the WS01 form • Provide training to help DWA assist and provide guidance others fill out WS01 form
Data checks and Quality Assurance	Documentation of ESR processes lacking	• ESR to improve documentation of their process - see option 2 above
	DWA data checking	• Provide improved guidance as to what should be checked and how • Add line to WS01 form for DWA to sign that information is checked and correct
	Data quality checks	• ESR to introduce ongoing data management checks to identify issues and report in addition to the data entry checks
	Quality assurance	• Ongoing investigations to evaluate data quality
Confidence in data	Lack of confidence in data	• Develop and publish data quality metrics possibly in the form of a data quality report published annually • Develop check list to ensure that data quality issues and taken into account when reporting on drinking-water quality data DQAC
Process	Lack of online registration and changes to Registration	• Improved data sharing/transfer between national and local versions of WINZ.
Data	Does WINZ hold all the data it should?	• Identify what other additional data needs users have

REFERNCES

Batini, C., Cappiello, C., Francalanci, C., & Maurino, A. (2009). Methodologies for Data Quality Assessment and Improvement. ACM Computing Surveys, , Vol. 41, No. 3, Article 16,.

Canadian Institute for Health Information. (2005). The CIHI Data Quality Framework. Ottawa, Canada: Canadian Institute for Health Information.

Coronial Services of New Zealand. (Undated). Data Quality. Retrieved January 30th, 2011, from Coronial Services of New Zealand: http://www.justice.govt.nz/courts/coroners-court/coronial-data-collection/data-quality

Ehling, M., Korner, T., & Nimmergut, A. (undated). Quality Indicators at the Federal Statistical Office Germany. Retrieved December 13, 2010, from www.oecd.org/dataoecd/0/23/18987057.ppt

ESR(2004). WINZ Advisory Group Meeting Minutes, Thursday 29 April 2004

Giovinazzo, W. (2009, August 18). BI: Only as Good as its Data Quality. Retrieved December 9, 2009, from Information Management: http://www.information-management.com/specialreports/2009_157/business_intelligence_bi_data_quality_governance_decision_making-10015888-1.html

Kerr, K. A., Norris, T., & Stockdale, R. (2008). The strategic management of data quality in healthcare. Health Informatics Journal , 259-266.

Kerr, K. (2006). The Development of a Data Quality Framework and Strategy for the New Zealand Ministry of Health. PhD thesis, University of Auckland .

London Health Services Centre. (2009, January 26th). Data Quality . Retrieved January 30th , 2011, from London Health Services Centre: http://www.lhsc.on.ca/About_Us/Case_Costing/Data_Quality.htm

Loshin, D. (2006). Monitoring Data Quality Performance Using Data Quality Metrics. Redwood City, CA: INFORMATICA.

McGilvray, D. (2008). Executing Data Quality Projects: Ten Steps to…. Sydney: Morgan Kaufman.

Ministry of Justice. (2008). Data Quality Framework.

Pipino, L. L., Lee, Y. W., & Yang, R. Y. (2002). Data Quality Assessment. Communications of the ACM , Vol 45, No 4, 211-218.

Wang, R. Y., & Strong, D. M. (1996). Beyond accuracy: What data quality means to data consumers. Journal of Management Information Systems , Vol 12,No 4, 5-34.

APPENDIX I: DATA QUALITY REVIEW PROJECT: FRAMEWORK

Introduction

The Data Quality Review Project examines how key water supply data are specified, gathered and used. At each step, from the initial requirements for information being defined to the eventual use of the information, assumptions are made that provide confidence in the integrity of the information. By examining these assumptions and the basis for making them, the consequences of the assumptions not being fulfilled can be identified.

The conceptual framework on which this project is based is presented in the figure below. The central column of the diagram sets out the steps associated with the use of data. The green column to its right identifies the assumptions that are made about each step in the process, and to the right risks that may result in the event of the assumptions not being met are identified. The "assumptions" and "risks" are what have been identified at the time of preparing the framework; it is possible that others may be identified as the project progresses.

To the left of the diagram the column, "Data handled" provides a generic description of what happens to the data and how it is used through the process. Examples of the legislation/regulations that specify requirements and the nature of the data that have to be gathered and assessed to establish whether these requirements are met are given in the left-most column.

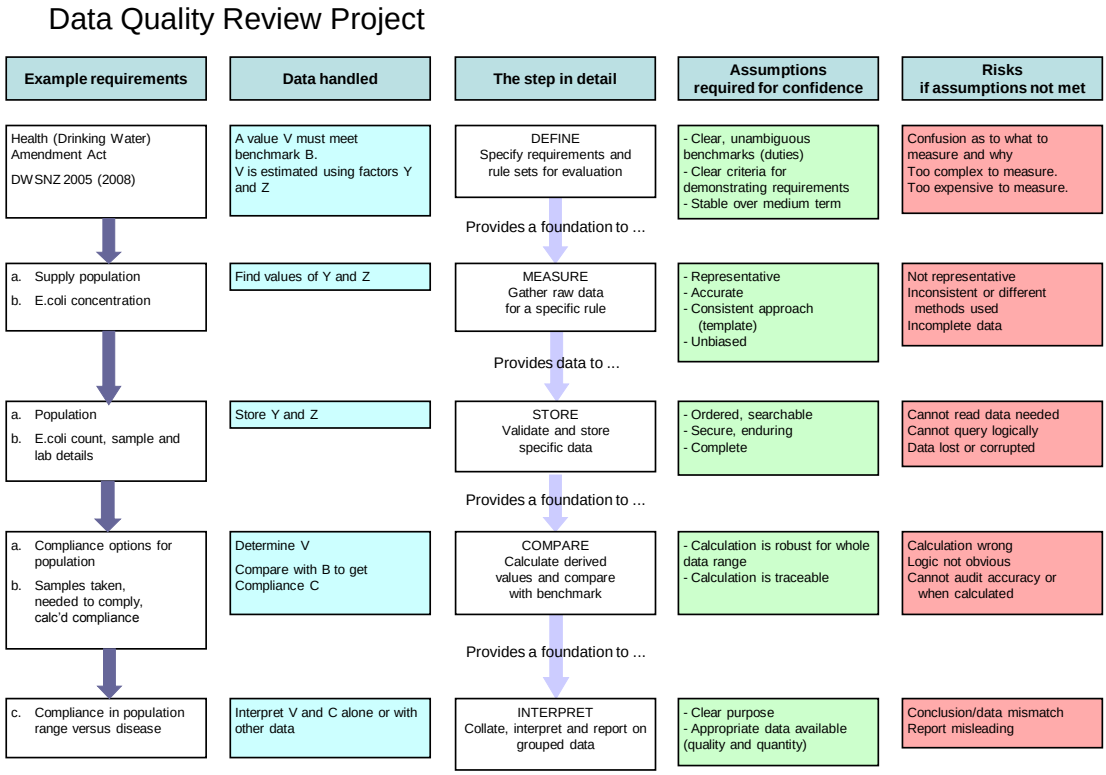

Figure 9: Process map for the data review project

APPENDIX II: SURVEY INTRODUCTION

The aim of the survey was to investigate drinking-water assessors' (DWAs) views on data and data quality. This was with specific reference to the *Register*, registration process, modifications to the register and general opinions about data quality held in WINZ.

Methods and Approach

An online survey was chosen as the most cost effective method to collect DWAs' views. A SurveyMonkey® PRO account was used to administer the survey.

The prime focus of the survey questionnaire was the *Register* and registration-related processes. This was decided upon during a workshop with ESR staff utilising a multicriteria approach. Semi-structured interviews followed with ESR staff and a DWA defined the scope and details of the questionnaires information requirements.

Using in-house expertise and following principles of best practice, the questionnaire was drafted and piloted. Those involved in the piloting stage included a DWA and a Ministry of Health representative.

The survey consisted of 36 questions with a combination of free text and tick-box responses. The questions were designed to collect a mixture of opinions and facts.

Prior to survey distribution a personalised e-mail was sent out on the 8 November 2010 to a PHU manager from each DHB, explaining what the survey was about. In addition, this email sought the PHU manager's permission for ESR to invite their staff to complete the survey. Positive responses were received from every DHB.

On the 16 November 2010, 31 individuals identified as DWAs or trainee DWAs, were sent an email inviting them to take part in the survey. A follow up e-mail was sent on the 29 November thanking those that had taken part in the survey and reminding those who had not that a response would be appreciated. The survey was closed on the 1 December, and some responses were received after this date.

The information collected by online survey was automatically entered into a database for analysis. There was one paper-based response which was entered manually.
The online approach was quicker than a postal survey as the raw survey results were immediately available for data analysis cutting out the data entry stage.

The survey outputs were analysed. Free text responses were analysed by coding the data and undertaking a form of content analysis to identify themes and issues. For tick box responses, descriptive statistics were produced and where appropriate bivariate and multivariate statistics analysis was undertaken. The resulting information was combined to produce an overview of the survey. This information was then combined with other lines of evidence in order to evaluate data quality. Late responses were included in the analysis, where practicable.

APPENDIX III: REQUEST TO PHU MANAGERS FOR STAFF TO TAKE PART IN SURVEY

Dear X,

Issue:

We are carrying out a project to evaluate the quality of data held within WINZ. With your permission, we would like to invite Drinking Water Assessor's from your DHB to take part in an online survey.

Required action:

Please could you respond within 3 days to let us know whether you give permission, or not, for ESR to approach your team to complete the survey. If you are not the appropriate person to provide permission, please could you advise who is.

Please feel free to forewarn your team that they will receive an e-mail from ESR inviting them to take part the survey.

Message:

As part of its contract with the Ministry of Health this year, ESR is undertaking a project to evaluate the accuracy, completeness, and fitness for purpose of the data held within WINZ. To do this we need to understand what happens to the data that are stored within WINZ from the very start of their journey i.e. the point at which the information is collected. We have a clear picture of which happens to the data once they reach ESR, but we need to be certain that we have an accurate understanding of the process through which the data pass before they reach us. We hope that the members of your team who collect drinking water quality data for WINZ would be able to assist us with this project.

To this end we have designed a questionnaire to discover how the raw information for WINZ is collected, and to understand your team's views on the information held within WINZ. The questionnaire will be administered through the web, a copy of which will be forwarded to you at the same time we contact your staff.

So that we can proceed with the work, could you please let us know whether you give permission for us to approach your staff. I would be grateful if you could provide your response to David.Wood@esr.cri.nz by Thursday 11 November.

Once your permission has been received we will contact your staff directly, based on the contact details within WINZ.

David Wood, has responsibility for the project and administering the questionnaire. Should you, or any of your team, have any questions, he can be reached by phone on 03 351 6019, or through email: David.Wood@esr.cri.nz.

Our thanks in anticipation of the help of you and your team.

Regards

David
David Wood
Senior Scientist
Water Programme
Environmental Science and Research Ltd
27 Creyke Rd, PO Box 29-181, Christchurch

APPENDIX IV: INVITITATION FOR DRINKING-WATER ASSESSORS TO TAKE PART IN SURVEY

Dear X,

ESR is currently undertaking a project to evaluate the drinking-water information it collects and uses on behalf of the MoH.

Analysis and interpretation of drinking-water quality data is needed to inform Ministry of Health policy, administer legislation and regulation, respond to questions from the Minister and parliament, and for local DHBs and water suppliers to take public health action. However, data analysis and interpretation are only as good as the accuracy, currency and completeness of the data. In turn, this quality and any constraints upon its usefulness or relevance need to be conveyed appropriately to the end users as part of the communication.

As part of this project we are asking for your help. Could you please take some time to complete this questionnaire. This will help us discover how the raw information for WINZ is collected, and to understand your views on the information held within WINZ. The information collected here will be used to perform an assessment of the current situation and consider improvements.

The information is being collected electronically using SurveyMonkey. To respond to this survey please click on this link:
http://www.surveymonkey.com/s/Data_Quality_Survey

Please could you respond by to the survey by 30th November.

All DHB's have given permission for their staff to take part in this survey and the questionnaire has been sent to all Drinking-Water Assessor's. We will not name any individuals in the survey write up, but as there is only a small number of Drinking-Water Assessor's it may be the case that you are identifiable by the comments you make. If you have any questions or concerns please contact David Wood either on: 03 351 0130 or David.Wood@esr.cri.nz

Thank you for your help
David

David Wood
Senior Scientist
Water Programme
Environmental Science and Research Ltd
27 Creyke Rd
PO Box 29-181
Christchurch

APPENDIX V: QUESTIONS AND QUATITATIVE RESULTS OF SURVEY

1. Introduction

ESR is currently undertaking a project to evaluate the drinking water information it collects and uses on behalf of the MoH.

Analysis and interpretation of drinking-water quality data is needed to inform Ministry of Health policy, administer legislation and regulation, respond to questions from the Minister and parliament, and for local DHBs and water suppliers to take public health action. However, data analysis and interpretation are only as good as the accuracy, currency and completeness of the data. In turn, this quality and any constraints upon its usefulness or relevance need to be conveyed appropriately to the end users as part of the communication.

As part of this project we are asking for your help. Could you please take some time to complete this questionnaire. The information collected here will be used to perform an assessment of the current situation and consider improvements.

The information is being collected electronically using SurveyMonkey.

The questionnaire has been sent to all DWA's. We will not name any individuals in the survey write up, but as there is only a small number of DWA's it may be the case that you are identifiable by the comments you make. If you have any questions or concerns please contact David Wood either on:

03 351 0130

or

David.Wood@esr.cri.nz

Data Quality Survey

SurveyMonkey

1. Do you use FORM: WS01 to apply for registration for a Water Supply?

	Response Percent	Response Count
Yes	95.8%	23
No	4.2%	1
answered question		24
skipped question		0

2. Does this form collect all the information that you and the water supplier need to provide to ESR in order to register a Water Supply?

	Response Percent	Response Count
Yes	83.3%	20
No	16.7%	4
Don't know	0.0%	0
answered question		24
skipped question		0

3. What additional information, if any, do you need to provide to ESR to register a supply?

	Response Count
	12
answered question	12
skipped question	12

4. Who is the most likely to complete FORM:WS01?

	Response Percent	Response Count
Water Supplier	**37.5%**	**9**
IANZ accredited DWA	33.3%	8
Trainee DWA	4.2%	1
HPO	4.2%	1
Other (please specify)	20.8%	5
	answered question	**24**
	skipped question	**0**

5. Are any checks performed on the accuracy of some or all of the information provided on FORM: WS01?"

	Response Percent	Response Count
Yes –all the information checked	**58.3%**	**14**
Yes –some of the information is checked	29.2%	7
No – the information is not checked	12.5%	3
	answered question	**24**
	skipped question	**0**

6. What are those checks, please could you describe them?

	Response Count
	19
	answered question	**19**
	skipped question	**5**

7. How often are you involved with registering a water supply?

	Response Percent	Response Count
Every week	4.3%	1
Monthly	0.0%	0
Few times a year	**95.7%**	**22**
Never	0.0%	0
Don't know	0.0%	0
	answered question	**23**
	skipped question	**1**

8. To what extent do you agree or disagree with the following statements?

	Strongly Agree	Agree	Neutral	Disagree	Strongly Disagree	Response Count
I find the registration process simple.	21.7% (5)	**56.5% (13)**	13.0% (3)	4.3% (1)	4.3% (1)	23
It is easy to provide the required information.	13.6% (3)	**40.9% (9)**	22.7% (5)	18.2% (4)	4.5% (1)	22
The registration process takes a long time.	4.5% (1)	4.5% (1)	22.7% (5)	**54.5% (12)**	13.6% (3)	22
The registration process could be improved.	9.1% (2)	22.7% (5)	**50.0% (11)**	13.6% (3)	4.5% (1)	22
					answered question	**23**
					skipped question	**1**

9. If you agreed that the registration process could be improved. Can you suggest how it could be improved?

	Response Count
	9
answered question	9
skipped question	15

10. Where do you get your definition of "Specified Self Supply" (please tick all that apply to you).

	Response Percent	Response Count
Standards	13.0%	3
Grading	0.0%	0
Legislation	**82.6%**	**19**
Tech manual	4.3%	1
Training	17.4%	4
Rulings (from the National Coordination Team or Ministry of Health)	69.6%	16
Colleagues	26.1%	6
Other (please specify)	13.0%	3
answered question		23
skipped question		1

11. Is the definition of 'specified self supply' clear to you?

	Response Percent	Response Count
Yes	52.2%	12
No	47.8%	11
answered question		23
skipped question		1

12. If you answered 'yes' to the previous question please describe what a specified self supply is? If you answered 'no' could you please describe what you consider to be unclear about the definition?

	Response Count
	22
answered question	22
skipped question	2

13. Who usually estimates the population figures (Please select one only)?

	Response Percent	Response Count
Water Supplier	91.3%	21
IANZ accredited DWA	0.0%	0
Trainee DWA	0.0%	0
HPO	0.0%	0
Other (please specify)	8.7%	2
answered question		23
skipped question		1

14. How do they usually estimate the population figures?	
	Response Count
	23
answered question	23
skipped question	1

15. For which of the following, if any, is determining the population most difficult? Please choose one.			
		Response Percent	Response Count
Network supply		8.7%	2
Bulk supply		8.7%	2
Prescribed supply		0.0%	0
Port/airport		26.1%	6
Specified self supply		26.1%	6
No opinion / No difference		**30.4%**	**7**
		answered question	**23**
		skipped question	**1**

16. What are your main sources of advice when you are estimating a population? Please tick all that are applicable

	Response Percent	Response Count
Standards	17.4%	4
Grading	8.7%	2
Legislation	17.4%	4
Tech manual	13.0%	3
Training	4.3%	1
Rulings	8.7%	2
Colleagues	21.7%	5
Not applicable, we do not help estimate population	**52.2%**	**12**
Others (please specify)	43.5%	10
	answered question	**23**
	skipped question	**1**

17. What do you see as the main difficulties, if any, in determining the population figures?

	Response Count
	16
answered question	**16**
skipped question	**8**

18. Do you provide GPS coordinates for all new registrations and modifications to existing registrations?

	Response Percent	Response Count
Yes	82.6%	19
No	17.4%	4
Not applicable	0.0%	0
answered question		23
skipped question		1

19. If no, what are the barriers to providing GPS coordinates? (please tick all that are relevant)

	Response Percent	Response Count
Lack of access to a GPS	0.0%	0
Takes too long to obtain the information	28.6%	2
It is up to the drinking water supplier to provide the info	0.0%	0
GPS coordinates have a low priority	0.0%	0
I don't know how to use a GPS	0.0%	0
Other (please specify)	85.7%	6
answered question		7
skipped question		17

20. What would make it easier to provide GPS coordinates?

	Response Count
	17
answered question	17
skipped question	7

21. How often are you involved with modifying the registering of a water supply?

	Response Percent	Response Count
Every week	0.0%	0
Monthly	8.7%	2
Few times a year	**82.6%**	**19**
Never	4.3%	1
Don't know	4.3%	1
	answered question	**23**
	skipped question	**1**

22. When a change occurs to a supply, there is a need to alter the register. How would you describe the state of the registrations in your area?

	Response Percent	Response Count
All up to date	4.3%	1
Mostly up to date	**78.3%**	**18**
Somewhat up to date	13.0%	3
Out of date	4.3%	1
	answered question	**23**
	skipped question	**1**

23. How do you usually find out that a change to registration needs to be made?

	Response Count
	23
answered question	**23**
skipped question	**1**

24. What is the usual reason for making a change?	
	Response Count
	22
answered question	22
skipped question	2

25. So, how important do you think these data items are?						
	Not at all important	Unimportant	Neutral	Important	Very important	Response Count
Name of supplier	0.0% (0)	0.0% (0)	0.0% (0)	13.0% (3)	**87.0% (20)**	23
Name of person responsible	0.0% (0)	0.0% (0)	8.7% (2)	30.4% (7)	**60.9% (14)**	23
Name of contact person	0.0% (0)	0.0% (0)	4.3% (1)	34.8% (8)	**60.9% (14)**	23
Supply type	0.0% (0)	0.0% (0)	4.3% (1)	21.7% (5)	**73.9% (17)**	23
Maximum daily volume	0.0% (0)	17.4% (4)	**43.5% (10)**	34.8% (8)	4.3% (1)	23
Community name	0.0% (0)	0.0% (0)	0.0% (0)	39.1% (9)	**60.9% (14)**	23
Community population	0.0% (0)	0.0% (0)	4.3% (1)	34.8% (8)	**60.9% (14)**	23
Seasonal changes in population	0.0% (0)	4.3% (1)	13.0% (3)	**47.8% (11)**	34.8% (8)	23
Treatment plant GPS coordinates	0.0% (0)	4.3% (1)	21.7% (5)	**47.8% (11)**	26.1% (6)	23
Treatment plant Location description	0.0% (0)	0.0% (0)	8.7% (2)	**60.9% (14)**	30.4% (7)	23
Source GPS coordinates	0.0% (0)	9.1% (2)	13.6% (3)	**45.5% (10)**	31.8% (7)	22
Source Location description	0.0% (0)	0.0% (0)	4.3% (1)	**56.5% (13)**	39.1% (9)	23
				answered question		23
				skipped question		1

26. To what extent do you agree or disagree with these statements

	Strongly disagree	Disagree	Neutral	Agree	Strongly agree	Response Count
The information held within WINZ is useful to me.	0.0% (0)	0.0% (0)	0.0% (0)	30.4% (7)	**69.6% (16)**	23
The information held within WINZ is useful to other DWAs.	0.0% (0)	4.3% (1)	0.0% (0)	43.5% (10)	**52.2% (12)**	23
The information held within WINZ is useful to Drinking Water Suppliers.	0.0% (0)	0.0% (0)	26.1% (6)	**39.1% (9)**	34.8% (8)	23
The information held by WINZ is useful to non-health professionals involved in supplying water such as engineers	0.0% (0)	4.3% (1)	**39.1% (9)**	**39.1% (9)**	17.4% (4)	23
The information held within WINZ is useful to other Health Professionals.	0.0% (0)	8.7% (2)	**43.5% (10)**	39.1% (9)	8.7% (2)	23
The information held within WINZ is useful to the MoH.	0.0% (0)	0.0% (0)	17.4% (4)	**43.5% (10)**	39.1% (9)	23
The information held within WINZ is useful to the public.	0.0% (0)	4.3% (1)	**52.2% (12)**	26.1% (6)	17.4% (4)	23
					answered question	**23**
					skipped question	**1**

27. How often do you use WINZ5 as part of your work?

	Response Percent	Response Count
Every week	**52.2%**	**12**
Monthly	39.1%	9
Few times a year	8.7%	2
Never	0.0%	0
Don't know	0.0%	0
	answered question	**23**
	skipped question	**1**

28. How often do you use WINZ6 as part of your work?

	Response Percent	Response Count
Every week	17.4%	4
Monthly	**52.2%**	**12**
Few times a year	30.4%	7
Never	0.0%	0
Don't know	0.0%	0
	answered question	**23**
	skipped question	**1**

29. Just think about the information in WINZ, to what extent do you agree or disagree with the following statements:

	Strongly disagree	Disagree	Neutral	Agree	Strongly agree	Response Count
Data held in WINZ is accurate- it correctly reflects the character of a water supply.	0.0% (0)	13.0% (3)	13.0% (3)	**52.2% (12)**	21.7% (5)	23
Data in WINZ is up to date.	0.0% (0)	17.4% (4)	26.1% (6)	**34.8% (8)**	21.7% (5)	23
Data in WINZ is comprehensive, it contains all the information it should have.	8.7% (2)	17.4% (4)	**34.8% (8)**	21.7% (5)	17.4% (4)	23
					answered question	**23**
					skipped question	**1**

30. Is data quality important to you?

	Response Percent	Response Count
Yes	100.0%	23
No	0.0%	0
answered question		23
skipped question		1

31. If data quality is not important to you, why not?

Response Count
2
answered question 2
skipped question 22

32. Are there any particular areas of concern, around data quality for you?

Response Count
16
answered question 16
skipped question 8

33. Are you a (please tick as appropriate):

	Response Percent	Response Count
IANZ accredited DWA	87.0%	20
Trainee DWA	8.7%	2
HPO	13.0%	3
Other (please specify)	4.3%	1
answered question		23
skipped question		1

34. Which Health Area do you currently work in?

35. How long have you been working in the drinking water area?

	Response Percent	Response Count
less than 1 year	0.0%	0
1-2 years	4.3%	1
2-4 years	34.8%	8
4-10 years	21.7%	5
10 years or longer	39.1%	9
	answered question	23
	skipped question	1

36. Thank you for completing this questionnaire, if you have any further comment about data and data quality or this survey, please write them below:

	Response Count
	9
answered question	9
skipped question	15

APPENDIX VI REPORT DISTRIBUTION

Copies have been made and distributed to:

Sally Gilbert (MoH)
David de Jager (MoH)

Viv Smith (ESR)
Jan Gregor (ESR)

Further copies of this report may be obtained from: ESR Christchurch Science Centre
P O Box 29 181
Christchurch