Government of **Western Australia**
Child and Adolescent Health Service

Clinical Audit Handbook

Sue Skull and Natasha Bear
Edited by Kylie Hill

Updated July 2020

This page has been intentionally left blank

Contents

Foreword

Clinical audits should be undertaken regularly by all health professionals. Audits entail a review of some aspect of clinical practice and represent an essential first step to identifying problems and inefficiencies for improvement. This Handbook is an opportunity to support audit processes for those needing training or a refresher in methodology. It is also useful for those planning to conduct other types of surveys. The Handbook is a deliberately generic and open access guide that is suitable for students and staff from any clinical discipline.

Key understandings underpinning this handbook are that:

1) An audit is a type of survey, or a cross-sectional study. All cross-sectional studies are a form of observational, descriptive research project. Data collection occurs via a questionnaire or data collection instrument (sometimes called an audit tool).

2) There are three potential approval pathways to consider when conducting a cross-sectional study in the health sector in Western Australia. The choice of pathway depends on a number of criteria, including consideration of ethical issues. These pathways are explained in the Handbook, and are the:

    - Governance Evidence Knowledge Outcomes pathway (used in public health settings) for quality assurance projects using low risk, routinely collected data. The data collected and reported are generally for internal use only. Most audits use this pathway.

    - Low and Negligible Risk pathway, for low risk data where specific study criteria apply. Participants generally provide consent.

    - Human Research Ethics Committee pathway, for projects where participants usually provide consent, data may be more sensitive or higher risk, and are generally intended for external use, such as scientific publication.

3) Three types of cross-sectional study designs, all related to service delivery improvement, are included in this Handbook are:
    - Baseline surveys.
    - Clinical audits comparing practice against an accepted standard.
    - Clinical audit cycles.

The Handbook can be used to either guide independent survey or audit completion, or as a resource for those undertaking components of these projects with variable levels of supervision. As a searchable document, it does not necessarily need to be read from front to back in its entirety, as it has been written such that each section can be understood without the need to read the preceding sections.

I am indebted to my co-author Natasha Bear, and the Perth Children's Hospital Foundation who generously supported our editor Kylie Hill, without whom the Handbook could not have been completed. Professor Donna Mak kindly shared her excellent University of Notre Dame Australia medical student Clinical Audit

Handbook, developed by herself, Dr Sally Murray and Dr Jelena Maticevic, and reviewed content. I would like to thank all our reviewers, listed below, for their helpful suggestions.

Handbook Contributors		
A/Prof Sue Skull	Head Research Education Program, Deputy Director Dept of Child Health Research, Perth Children's Hospital, WA	Conception Program lead Lead author
Dr Natasha Bear	Biostatistician, Dept of Child Health Research, Perth Children's Hospital, WA	Biostatistical author
Dr Kylie Hill	Medical Editor	Formatting and editing text
Ms Alison Hopkins	Promotions and Communications Officer, Dept of Child Health Research, Perth Children's Hospital, WA	Administration and web maintenance
Dr Christina Bertilone	Senior Medical Advisor, Patient Safety and Clinical Quality Directorate, Clinical Excellence Division, Dept of Health WA	Reviewer
Ms Marnie Chellew-Hawley	Clinical Governance Unit, Perth Children's Hospital, Perth, WA	Reviewer
Dr Matthew Cooper	Manager of Biometrics, Telethon Kids Institute, Perth, WA	Reviewer
Dr Darren Gibson	Director of Research, Dept of Health WA	Reviewer
Ms Sonia Harwood	Clinical Governance Unit, Perth Children's Hospital, Perth, WA	Reviewer
Ms Jaime Henderson	Safety and Quality, Child and Adolescent Health Service, Dept of Health WA	Reviewer
Prof Donna Mak	Professor and Head of Population and Preventive Health, School of Medicine, University of Notre Dame, Perth, WA	Reviewer
Dr Sarah Trinder	General Paediatrics Advanced Trainee, Perth Children's Hospital, WA	Reviewer
Sean Howarth	Ethics Coordinator, Sir Charles Gairdner Hospital	Reviewer

The Handbook is web-based and date marked on the front cover according to the last update.
We welcome feedback at any time on the Handbook via ResearchEducationProgram@health.wa.gov.au

A/Prof Sue Skull
Head, CAHS Research Education Program
Deputy Director, Department of Child Health Research
Perth Children's Hospital, Western Australia

List of abbreviations

CAHS	Child and Adolescent Health Service
CI	Confidence Interval
GEKO	Governance Evidence Knowledge Outcomes
HREC	Human Research Ethics Committee
LNR	Low and Negligible Risk
NHMRC	National Health and Medical Research Council
PCH	Perth Children's Hospital
QA	Quality Assurance
QI	Quality Improvement
WA	Western Australia

Resources and references embedded throughout this Handbook

A number of resources and references have been embedded throughout this Handbook. Where information has been extracted directly, the source document is cited at the end of the relevant paragraph. At the end of each section, a section called 'additional resources' has been added, which includes weblinks and publications grouped according to topic. The reader is encouraged to review this list. Section 9 is the reference list for sources that have been cited in the Handbook. Finally, where appropriate, the Handbook cross-references other sections that provide related information.

Section 1: Background

This section defines the terms survey and audit and also describes the ethical considerations related to undertaking these types of project.

1.1 Definitions

Survey

A survey is a type of cross-sectional (also known as observational or descriptive study) where data are collected during a 'snapshot' in time. ***An audit is a type of survey.*** Surveys can be done as part of a quality assurance (QA) activity, or as research requiring ethics approval.

Baseline survey

A baseline survey occurs where investigators describe an issue(s) or practice before standards for performance have been set. That is, they 'capture' details related to current practice or issues.

Clinical audit

A clinical audit is a survey which involves a comparison of actual practice or activity against established or accepted standards of practice. The aim is to improve clinical care and outcomes through a systematic review of practice against explicit criteria such as a clinical practice guideline or benchmark (adapted from National Institute for Clinical Excellence). In healthcare settings, clinical audits are part of clinical governance, which aims to ensure that patients receive the best quality care (Benjamin, 2008). A clinical audit is not, for example, simply a "stock take" or description of characteristics of a group of people with a certain condition.

Clinical audit cycle

Following a clinical audit and implementation of change, further monitoring is used to confirm that there has been an improvement in healthcare delivery. An audit that is repeated (multiple times depending on requirements), as part of a (continuous) quality improvement (QI) process is referred to as a clinical audit cycle (Figure 1) (Benjamin, 2008).

Note: The National Health and Medical Research Council (NHMRC) defines quality assurance activities as those 'where the primary purpose is to monitor or improve the quality of service delivered by an individual or an organisation.' Terms such as 'peer review', 'quality assurance', 'quality improvement', 'quality activities', 'quality studies' and 'audit' are often used interchangeably.

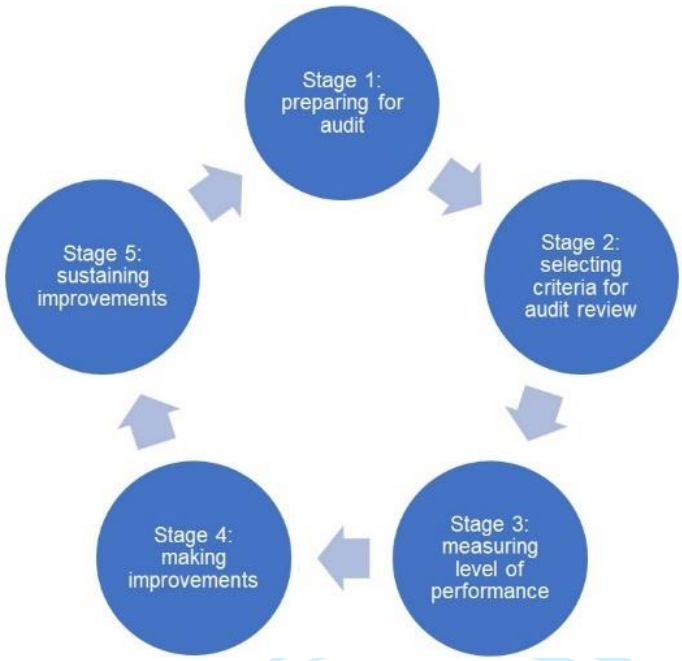


Figure 1: The clinical audit cycle

Weblinks to source information

Ethical considerations in quality assurance and evaluation activities (NHMRC, Australia, 2014)
https://nhmrc.gov.au/about-us/publications/ethical-considerations-quality-assurance-and-evaluation-activities

Principles for best practice in clinical audit (National Institute for Clinical Excellence, National Health Service, UK, 2002)
http://www.uhbristol.nhs.uk/files/nhs-ubht/best_practice_clinical_audit.pdf

1.2 What can cross-sectional studies tell us?

As cross-sectional studies such as surveys and audits record variables (outcomes and exposures) during a 'snapshot' in time, they are the most appropriate study design to choose when the aim is to:

- Compare current practice against an expected standard. Example study questions include:
    - To what extent are hospital antibiotic prescribing guidelines being followed?
    - What proportion of children admitted to a tertiary paediatric hospital have vaccination status recorded in their hospital file?

- Describe a clinical issue such as the burden of a disease (e.g. prevalence, severity, cost) and form a basis for benchmarking against the published literature, established national standards or other settings. Example study questions include:
    - What proportion of children and adolescents presenting to the Emergency Department had their vaccination status checked?
    - What proportion of these children and adolescents required vaccination?
    - Did those children and adolescents who required vaccination, differ by age or gender or postcode?
- Describe some aspect of clinical practice. Example study questions include:
    - How many times are intubated and ventilated patients in the intensive care unit suctioned over a 24-hr period?
    - What is the median wait time in the Emergency Department before a patient is seen by a doctor?
- Develop hypotheses for future study.

Important limitation: Cross-sectional studies essentially 'observe without interference' at a point in time, and the data obtained can be used to generate hypotheses, but not test relationships between exposures and outcomes. That is, a survey cannot definitively answer questions like 'what are the causes for this outcome', or 'does this practice work', or 'how well does it work'. This is because, with a simple cross-sectional study design, all variables are measured at one point in time and it is not possible to account for the many factors that may have contributed to the outcome of interest, including natural trends over time. As cross-sectional studies are unable to definitively link cause (or exposure) with effect (or outcome), they are 'lower down' the hierarchy of study design (Figure 2). This limitation is important to remember when interpreting and reporting audit results (see Section 6 for more information on interpreting audit results).

However, surveys and audits remain a very important and useful study design and can generate a hypothesis for causality that could then be examined by a 'higher order' study design, such as a case-control or cohort study, or randomised controlled trial. For example:

- As part of a clinical audit, it is noted that more than 50% of children and adolescents with asthma who are seen by the respiratory medicine clinic are overweight or obese. Asthma control seems worse in those who are overweight or obese. As these data were collected as part of a cross-sectional study, they cannot demonstrate whether worse asthma control is a risk factor for the increased weight, or the increased weight is a risk factor for worse asthma control. Based on this, a hypothesis is developed that in children and adolescents with asthma, weight loss will improve asthma control. This hypothesis could then be tested with a future cohort study or nested case-control study.

Figure 2: The hierarchy of evidence obtained using different study designs

1.3 Why do an audit?

As in any workplace, errors in practice, suboptimal practice, or inefficiencies can occur in our healthcare system, despite the training and best intentions of healthcare professionals. Detection of problems is greatly improved when audits of practice are undertaken. Audits examine how well clinical care is being provided and help to identify opportunities for improvement. Audits are ideally part of a continuous process for facilitating quality improvement.

Some of the benefits of undertaking a clinical audit are to:

- Assess and improve quality and efficiency of patient care.
- Uphold professional standards/facilitate professional development.
- Identify and measure areas of risk.
- Create a culture of transparency and quality improvement, where regular audit activity occurs.
- Educate health professionals and keep up to date with evidence-based practice.

What makes a suitable topic?

For a clinical audit, the topic must measure some aspect of practice against a benchmark or standard. It should not aim to find new information about a clinical condition or practice, as this would not meet audit criteria. The topic should have potential to improve practice or policy. Audits are ideally multidisciplinary. It is best to choose audit topics based on:

- High risk, high volume, or high cost problems,
- National clinical audit or service frameworks; or
- Institution-based set priorities.

In addition, think about the time and resources available for, and required to complete the project. For large audits, a multidisciplinary team may be required over an extended period. Where the scope is large and the outcomes have major implications for future practice across a large sector, it is important to include all relevant stakeholders from the outset, including community members. An inclusive approach is far more likely to result in a meaningful topic, appropriate design, successful conduct and ultimately result in the implementation of a change in policy and/or practice.

It is important to be realistic. For example, if a student is available for just a three-month time commitment in a part-time capacity, it may be appropriate for them to be responsible for just the literature review, or the design of the audit, or in the case of an existing clinical audit cycle, perhaps just the next data collection or reporting phase. All are important contributions worthy of inclusion as an author in any reports arising from the completed project. Consideration should be given to availability, stage of development, required sample size and resources available.

Weblinks to source information

A practical guide to clinical audit (Clinical Governance Support Team, National Health Service, UK, 2005)
https://webarchive.nationalarchives.gov.uk/20081112120728/http://www.cgsupport.nhs.uk/Resources/Clinical_Audit/1@Introduction_and_Contents.asp

A practical guide to clinical audit (Quality & Patient Safety Directorate, Ireland, 2017)
https://www.hse.ie/eng/about/who/qid/measurementquality/clinical-audit/practicalguideclaudit2013.pdf

1.4 Is ethical approval required?

'But this is just an audit!'

Quality assurance, evaluation and research exist on a continuum of activity, and work that begins as one form of activity can evolve into another over time. In addition, 'irrespective of whether an activity is QA, evaluation or research, the activity

must be conducted in a way that is ethical. This should include consideration of whether the people involved will be exposed to any harm as a result of the activity. Those conducting the activity need to consider a range of issues including consent, privacy, data security, relevant legislation, national/professional standards and whether ethical review is required' (https://nhmrc.gov.au/about-us/publications/ethical-considerations-quality-assurance-and-evaluation-activities). Possible harms may include physical, emotional or financial (and more). Intention to publish in a peer-reviewed scientific journal is generally a trigger for higher level review.

Even if a survey/audit is being conducted as a QA activity for in-house use only, it will always require consideration of ethical issues. As the scope of a project is refined, sometimes it will require approval from a Human Research Ethics Committee (HREC) or consideration of a Low and Negligible Risk pathway depending on the data collected, methods used and intention to publish or present outside the institution where data were collected. Where personal data are to be included, it is important to consider 'whether the people involved (e.g. participants, staff or the community) will be exposed to any risk, burden, inconvenience or possible breach of their privacy' (https://nhmrc.gov.au/about-us/publications/ethical-considerations-quality-assurance-and-evaluation-activities).

For audits, 'the data being collected and analysed are expressly for the purpose of maintaining standards or identifying areas for improvement in the environment from which the data were obtained.' In addition, 'there must be no triggers for ethical review' (https://nhmrc.gov.au/about-us/publications/ethical-considerations-quality-assurance-and-evaluation-activities). Audit data are generally intended for internal use. If data are to be shared outside the setting in which they were collected (e.g. at a conference, or for publication, where data can be easily picked up and distributed more widely), then ethical review may be required. ***Data collected as part of an audit may still be sensitive, and the broader ethical concerns must be considered.*** Examples of this are presented below:

Example 1: An audit is conducted of critical incidents that occurred in a hospital over a two-year period, based on interviews with staff about whether these were reported or acted upon. As many as 20% of critical incidents were found to have gone unreported based on data collected in this sample. A student reports the results at a local conference with practical suggestions for improvement. Personal phones were used by those in attendance to photograph the data slides, including reporters from the city's main newspaper, who find this interesting and decide to run a story. This appears the next day on the front page of the main newspaper under the heading: 'Our doctors and nurses covering up mistakes.' Furthermore, a specific example is given of an easily identifiable staff member based on an 'off the cuff' remark during question time.

Example 2: An audit of a clinic that predominantly treats refugees reports on their routinely investigated infectious diseases. An article is subsequently published that includes rates of certain infections that are much higher than the general population, listed by country of origin. Local refugee leaders are outraged as they feel they have been labelled as a 'problem' for the wider community and will be persecuted.

The following is a list provided by NHMRC (https://nhmrc.gov.au/about-us/publications/ethical-considerations-quality-assurance-and-evaluation-activities) of important factors that may trigger the need for ethical review:

- 'Where the activity potentially infringes the privacy or professional reputation of participants, providers or organisations.
- Secondary use of data - using data or analysis from QA or evaluation activities for another purpose.
- Gathering information about the participant beyond that which is collected routinely. Information may include biospecimens or additional investigations.
- Testing of non-standard (innovative) protocols or equipment.
- Comparison of cohorts.
- Randomisation or the use of control groups or placebos.
- Targeted analysis of data involving minority/vulnerable groups whose data is to be separated out of that data collected or analysed as part of the main QA/evaluation activity.'

Where one or more of the triggers above apply, the guidance provided in the National Statement on Ethical Conduct in Human Research, 2008 (updated 2018) should be followed. Ethical review can occur at different levels depending on the nature of the project and the data collected.

What are the options for ethical review and approval?

There are three potential approval pathways to consider when conducting a cross-sectional study (survey or audit) in the health sector in Western Australia (WA). These options conform with the NHMRC statement, which states that ethical review can be undertaken at various levels, according to the level of risk involved in the research.

Below are three examples of audit topics that are related to antibiotic stewardship, but require different approval pathways:

- An audit of the proportion of prescriptions of antibiotic X that were prescribed in accordance with hospital antibiotic prescribing guidelines; this audit would not require HREC approval, but should be registered with an audit approval pathway such as the Governance Evidence Knowledge Outcomes [GEKO] system (described below).
- The above audit is expanded to include collecting data on medical staff's knowledge of, and attitudes to, hospital antibiotic prescribing guidelines; this study would probably be eligible for the LNR pathway.

- The above study was further expanded to include what is happening in another health service, e.g. antibiotics prescribed by the general practitioner before presenting to hospital X, or exploring health consumer's (including patients/parents attending hospital X and people who have not attended hospital X) knowledge of, and attitudes to antibiotic stewardship; this study will need full HREC review and approval.

Where there is uncertainty about the best approval pathway for a project, discussion with staff from the local Clinical Governance and/or Safety and Quality Unit, Research Governance or Ethics Office is appropriate. Getting advice early can avoid problems later on, such as being unable to obtain retrospective approval at the correct level to publish project data.

No formal HREC review

Projects not requiring formal HREC review still require consideration of ethical issues. In the WA Public Health sector the mechanism for review of these projects, ***provided they are also QA/QI activities***, is the GEKO system. GEKO is accessible to all WA Health staff via their Health Employee log in. It is an on-line system that facilitates rapid on-line review by Committees generally managed by the Safety and Quality departments of most major WA Health hospitals and public health sites. Audit reports and outcomes are also recorded within GEKO. Each health service uses different processes for review of activities including publication.

If requested and approved, a publishing reference number is sent to the Investigator. Guiding documents for this review are the NHMRC Ethical Considerations in Quality Assurance and Evaluation Activities and National Statement on Ethical Conduct in Human Research (see weblinks to source information below).

To enable the Committee and in some cases Executive signatory to assess each quality assurance/improvement project as suitable for the GEKO pathway and/or for publication, it is important to outline any potential ethical issues in the application and to provide the data collection instrument for consideration. Populations that are excluded under the LNR pathway would normally also be ineligible for submission to the GEKO pathway. However, the NHMRC guidelines provide scope for flexibility depending on the nature of the project. Some of these projects may need HREC review.

Although different sites may use different processes, health professionals conducting quality assurance/improvement activities outside WA Health are still bound to comply with the national standards set down by NHMRC. These state: 'irrespective of whether an activity is QA, evaluation or research, the activity must be conducted in a way that is ethical. This should include consideration of whether the people involved will be exposed to any harm as a result of the activity. Those conducting the activity need to consider a range of issues including consent, privacy, relevant legislation, national/professional standards and whether ethical review is required.' (https://nhmrc.gov.au/about-us/publications/ethical-considerations-quality-assurance-and-evaluation-activities)

Examples of projects that do not require a formal HREC review include:

- A study of patients undergoing colorectal surgery at a tertiary hospital examining what proportion have received antibiotic prophylaxis in accordance with established standards of practice outlined by the Therapeutic Guidelines Antibiotic Expert Group
- A study of diabetic patients who are currently on a general practitioner management plan examining what proportion have undergone diabetic retinopathy screening in accordance with the Royal Australian College of General Practitioner's recommendations for diabetic care.

Low and negligible risk pathway

In WA, most but not all public health sites with HRECs will also have an LNR pathway as an alternative to the full review process undertaken by an HREC. The LNR pathway is an 'express' process with ethical review and is usually considerably faster than a full HREC review (e.g. three weeks). In general, this is for projects where data collected are considered to be low risk. Risk assessment needs to consider a range of impacts including physical, psychological, emotional, financial and/or legal harms.

The NHMRC defines research as 'low risk' where the only foreseeable risk to the participant is one of discomfort,and 'negligible risk' where the only foreseeable risk to the participant is one of inconvenience. Studies that may be suitable for the LNR pathway include most qualitative research projects, and any study where the data is collected by questionnaire and/or focus groups and the target population is not excluded by the criteria set out by the NHMRC. Studies that describe any of the following features are ***unlikely to be eligible*** for the LNR pathway:

- Any potential risk to the participant which will cause them anything more than discomfort.
- An intervention: for example, the use of drugs or devices; taking specimens from children and public and mental health interventions that would cause the participant anything more than discomfort.
- Vulnerable individuals: for example, people who have a dependent relationship with medical personnel, mental illness, cognitive or intellectual impairment, gender identity issues, or who are involved in illegal activities (e.g. illicit drug use).
- Aboriginal people or Torres Strait Islanders as the target study population.
- Genetic testing.
- Stem cells or their products.
- The creation of a databank, biobank or registry.
- The examination of sensitive personal or cultural issues.
- Women who are pregnant or their foetuses either in utero or ex utero.

- A request for either a ‘Waiver of Consent’ or permission to ‘Opt-out of Consent.’

Examples of projects suitable for the LNR pathway may include:

- A grounded theory (qualitative) study which explores the experiences of nurses who encounter young people with mental health problems; data collection will involve adult health care providers completing a questionnaire after obtaining consent.

- A qualitative exploration of the experiences and needs of parents of a child diagnosed with Type 1 diabetes when one parent has Type 1 diabetes; data collection will occur by voluntary completion of a questionnaire after obtaining consent.

Full HREC review

All other projects require review by a HREC and associated Governance offices where appropriate. Of note, audits across multiple sites, or that access state-wide data collections held by the Department of Health will require HREC review. Where there is uncertainty about the best approval pathway for a project, discussion with staff from the local Clinical Governance and/or Safety and Quality units, Research Governance or Ethics Offices is appropriate. Key reference documents are listed below. See Section 1.6 for information on the HRECs in WA.

Examples of projects that require review by a HREC may include:

- A prevalence study of type 2 diabetes amongst obese young adults involving collection of new data from non-routine questionnaires and blood tests.

- A prevalence study of risk-taking behaviours in adolescents collected prospectively via questionnaire.

Weblinks to source information

Ethical considerations in quality assurance and evaluation activities (NHMRC, Australia, 2014)
https://nhmrc.gov.au/about-us/publications/ethical-considerations-quality-assurance-and-evaluation-activities

National Statement on Ethical Conduct in Human Research (NHMRC, Australia, 2007 - Updated 2018)
https://nhmrc.gov.au/about-us/publications/national-statement-ethical-conduct-human-research-2007-updated-2018#block-views-block-file-attachments-content-block-1

1.5 Overview of steps required for a baseline survey or clinical audit

Conducting a baseline survey or an audit will require careful planning and stakeholder engagement. This will be covered in detail in Sections 2 to 7 but is outlined here in brief. The more effort that goes into the 'front end' planning of a project, the greater its chance of successful completion and meaningful outcomes with impact.

- Develop a study question and define project scope.
- Select a standard or benchmark (for audits).
- Determine the key stakeholders.
- Determine the audience for reporting.
- Develop a knowledge translation plan.
- Choose the study design (baseline survey, clinical audit or clinical audit cycle).
- Choose the sampling method.
- Choose the method for survey delivery.
- Maximise the response rate.
- Determine the sample size.
- Plan the analyses and choose variables for data collection.
- Design the data collection instrument.
- Set up the database.
- Pilot the data collection instrument and delivery method and train study personnel.
- Determine the approval pathway and obtain appropriate approvals.
- Collect and enter the data.
- Clean the database.
- Analyse the data.
- Critically interpret the results.
- Report the results / ensure knowledge translation occurs.
- Implement the recommendations.

1.6 Additional resources

Cochrane review on the effect of audits on clinical practice
Ivers, N., Jamtvedt, G., Flottorp, S., Young, J.M., Odgaard-Jensen, J., French, S.D., O'Brien, M.A., Johansen, M., Grimshaw, J., Oxman, A.D. (2012). Audit and feedback: effects on professional practice and healthcare outcomes. *Cochrane Database Syst Rev*. 13;(6):CD000259

Research Governance Framework

Research Governance Framework (Government of WA, Department of Health)
https://rgs.health.wa.gov.au/Pages/Research-Governance-Framework.aspx

Ethics information to undertake data linkage

https://www.datalinkage-wa.org.au/privacy-and-ethics/ethics/

Weblinks to CAHS Research Education Program content

Research Fundamentals (CAHS, Australia, 2019)
https://cahs.health.wa.gov.au/Research/For-researchers/Research-Education-Program/Past-seminars

Ethics Processes within WA Health (CAHS, Australia, 2019) which includes a list of HRECS within WA
https://cahs.health.wa.gov.au/Research/For-researchers

Section 2: Initial planning

This section provides a comprehensive overview of the methodological considerations associated with undertaking an audit or survey.

2.1 Develop a study question and define project scope

Following a review of the literature and discussion with peers, a clear, answerable and worthwhile question must be stated, along with aims and objectives, and rationale for undertaking the project (Aslam, 2010). This includes ensuring there is a good fit with stakeholder needs and that there will be impact potentially on policy and practice.

A project ***aim*** is a general statement about the intended outcome or goal. For example, 'are we meeting best practice standards for opportunistic vaccination?' A project aim is then generally broken down into a number of smaller objectives.

A project ***objective*** is a step with measurable outcomes that describes how an aim or goal has been achieved. Objectives should be highly focused and precisely described. A useful framework for each objective is the PICOT structure used for study questions (participant, intervention, comparator, outcome, time frame). For an audit or survey there will not be the intervention or comparator components. For example:

- Of all patients attending clinic X, what proportion have their vaccination status recorded for the period April to September 2019?

- Of all patients from clinic X with an incomplete vaccination status according to the Australian Immunisation Register, what proportion received an opportunistic vaccination at the time, for the period April to September 2019?

- Of all patients from clinic X with an incomplete vaccination status according to the Australian Immunisation Register, what proportion had a follow up appointment provided for their vaccination, for the period April to September 2019?

Clear objectives then facilitate identification of a standard or benchmark for comparison and inform development of the data collection instrument for the project.

Weblinks to source information

Asking focused questions (Centre for Evidence-Based Medicine, University of Oxford, UK)
http://www.cebm.net/index.aspx?o=1036

PICO: Formulate an answerable question (Cochrane Collaboration)
https://ph.cochrane.org/sites/ph.cochrane.org/files/public/uploads/Unit_Five.pdf

2.2 Determine the key stakeholders

Early and ongoing engagement with key stakeholders will ensure a useful question is asked, the right data are collected, and a communication and knowledge translation plan with impact is developed. Stakeholders are any individuals or groups who may have an interest in or be affected by the conduct or results of an audit (Mak, 2019). For example, stakeholders may include patients or their families, community representatives, clinical staff, or policy makers. Whereas an inclusive approach will enrich an audit with broader experience, knowledge and ideas, excluding key stakeholders can mean the difference between successfully conducting an audit and meeting resistance and failing, or between translating results into practice change and having them rejected.

2.3 Determine the audience for reporting

Survey data may be presented via publication, presentation at conferences, community or participant workshops or bulletins, media such as radio, television, newspaper or social media, to policy makers and others. Generally, audit data are for 'in-house' use only. Ethical review must always be considered where data are to be disseminated outside the organisation where they were collected, particularly if sensitive, or if the population for study is easily identifiable.

2.4 Develop a knowledge translation plan

The NHMRC states 'The creation of knowledge does not, of itself, lead to widespread implementation and positive impacts on health. The knowledge must be translated into changes in practice and policy for the benefits to flow to Australians. It can also take many years for research evidence to reach clinical practice.' It is therefore important at an early stage in project development to include a knowledge translation plan, particularly where the aim is to monitor and improve practice. Given the aim for all projects is for data to have maximum impact and reach all relevant stakeholders as soon as possible following completion, it is key to create ***and then follow*** a knowledge translation plan. Stakeholders always include participants, the broader community from which they are drawn, the scientific community, and staff involved with different aspects of patient care, but there are many others to consider such as government, policy makers and so on. The plan also helps determine the correct approval pathway; for example, a plan for external publication may trigger the need for HREC approval.

Some potential activities to plan for and document related to the impact of an audit or survey may include the following (see Section 7 for more information / examples of knowledge translation plans):

- Policy documents or briefs.
- Government/health organisation reports or clinical practice guidelines.
- Publications.
- Citations or commentaries related to publications and reports.

- Requests for permission to adapt methods or findings in other settings.
- Media reports or news about the project.
- Survey instruments developed.
- New positions created.
- Book chapters informed by the project data.
- Training workshops provided to knowledge users.
- Invitations to meetings to disseminate results to inform users, policy, education, practice.
- Invitations to advisory committees, boards, regulatory committees.
- Memberships and/or contributions to clinical practice guideline development teams.
- Consultancies.
- Training delivered/contracted for health professionals, patients, and/or policy makers.
- Award(s) or other formal recognition for research translation activities.
- Summary statements of indicators of impact.
- Indicators of impact from the research for graduate students.
- Elected membership on a society for which membership requires demonstration of project impact.
- Stories of impact from knowledge users (e.g., changes in practice/processes locally, or at a state or national level.
- Integration of project or program materials into services within a community.
- Contracts from health services for projects focused on changing practice.
- Knowledge users indicated on publications or grants.
- Course content or curriculum informed by the project.
- Social media use (e.g. Twitter [re-tweets], Linkedin or Facebook page downloads).

Weblink to source information

The Research Education Program: knowledge translation (CAHS, Australia, 2019) https://cahs.health.wa.gov.au/Research/For-researchers/Research-Education-Program/Past-seminars

2.5 Choose the study design

Baseline surveys and clinical audits are most often conducted via one-off administration of a questionnaire or survey instrument (for audits, these are sometimes referred to as an audit tool). Questions evaluate processes and procedures for the care of patients or performance of a service. Examples include:

- Establishing baseline activity of an aspect of clinical practice (baseline survey).
- Assessing how well an aspect of clinical practice is performing against an expected standard (clinical audit).

Select a standard or benchmark

For a ***clinical audit***, a 'standard' is required for comparison against measured performance. This may be a statement of best practice, or of the quality of care to be achieved, or a target for expected adherence. The expected level of adherence needs to be defined before commencing an audit and is often expressed as a percentage. Choosing a standard allows the clinical audit question to be framed in terms of a justifiable and measurable outcome, as well as helping to determine key variables for inclusion in the data collection instrument.

Example 1: Hospital policy states that patients arriving at the Emergency Department are to be reviewed and either admitted, discharged or transferred within four hours from the time of triage. For this audit, stakeholders decide to set the adherence level at 80%. The audit target then becomes:

- 80% of patients arriving at the Emergency Department are to be reviewed and either admitted, discharged or transferred within four hours from the time of triage.

Example 2: If 100% of hospital staff are required to have influenza vaccination annually, then you would expect '100% of hospital staff to have received influenza vaccine according to the hospital guidelines.' A comparison can then be made between measured and expected levels. If the national average for hospital staff influenza vaccination is actually lower than this, then it may be reasonable to also have a statement that you would expect 'the proportion of staff that have received annual influenza vaccination to be no lower than the national average.'

A standard should be chosen based on a thorough literature review and careful examination of the level of evidence for it, and then justified in subsequent reports.

Individual institutions or sites may set their own standards, but they may also exist at state and national level. For example, the **National Safety and Quality Health Service Standards** are developed by the ***Australian Commission on Safety and Quality in Health Care (ACSQHC)*** as part of the ***Australian Health Service Safety and Quality Accreditation Scheme***. They are designed to deliver a consistent level of safe and high-quality care across health services in Australia. The Commission

also sets Clinical Care Standards for various clinical domains, and has produced indicator specifications that allow for standardised data collection for the purposes of quality improvement.

There are times when a standard is not available. Some service evaluations collect baseline information to inform current practice. The systematic assessment of current practice, without comparison against set criteria, is a ***baseline survey***, also known as a baseline assessment practice review.

An ***audit cycle*** involves repetition of an audit following implementation of change after a baseline audit has been conducted, and time for changes to occur. The audit cycle can be conducted as a simple 'before-after' design, or a more complex time series design. The approach chosen has implications for analysis and reporting.

Weblinks to source information

How to use the evidence (NHMRC, Australia)
https://www.nhmrc.gov.au/about-us/publications/how-use-evidence

National Safety and Quality Health Service (NSQHS) Standards (Australia)
https://www.safetyandquality.gov.au/our-work/assessment-to-the-nsqhs-standards/
https://www.nationalstandards.safetyandquality.gov.au/

Clinical Care Standards (Australian Commission on Safety and Quality in Health Care, Australia, 2019)
https://www.safetyandquality.gov.au/standards/clinical-care-standards

Before-after audit cycle design

A before-after audit cycle design involves collecting data at two time points; once before the implementation of a change and once after the implementation of a change. The change could be the introduction of a policy or program designed to improve current practice. This cyclic process can be repeated multiple times.

Survey instruments must be designed with repetition in mind. For example, where the survey is to be repeated at different times of year, 'in the last three months' would be preferable to 'for the months October to December inclusive'.

For example, an initial audit indicates that adherence with hand hygiene is 70%. Following a hand hygiene awareness and training program, when re-measured, adherence with hand hygiene guidelines is 85%.

The before-after method does not allow for existing time trends to be accounted for and is therefore a less strong design than a time series design (below). As for all audits, it also cannot demonstrate that the training program was responsible for the improvement. This is because a simple survey cannot adjust for other factors that may also have influenced the outcome between the two time points (e.g. changes in staff, policy, funding, equipment, other training exposure, etc). The results must

therefore be reported objectively with this in mind (see Section 6 for more information on interpretation).

Time series design

A time series design can involve collecting data at multiple times before and after a practice or policy change. It differs from a simple 'before-after' design as the multiple measurements allow for natural time trends to be considered. It is therefore considered a stronger design compared to a simple 'before-after' approach (Harris, 2006). It is a popular method in quality improvement as it can measure improvement over time without the need for a control group (Andersson Hagiwara, 2016). Although more robust, it remains a cross-sectional design and still cannot definitively adjust for confounders or prove cause and effect. It therefore remains a hypothesis-generating study design. For example, Figure 3 demonstrates the change in hand hygiene adherence following a hand hygiene awareness and training program. There are 10 time points at which data were collected for the first audit, followed by 10 time points during the second audit. There are specific data analysis techniques to accompany this design. Commonly used methods include statistical process control and interrupted time series with segmented regression. See Section 5 for more information on statistical tests. Also, see Section 9 (References) for details of the following relevant readings Andersson Hagiwara (2016), Bernal (2017), Fretheim (2015) and Penfold (2013).

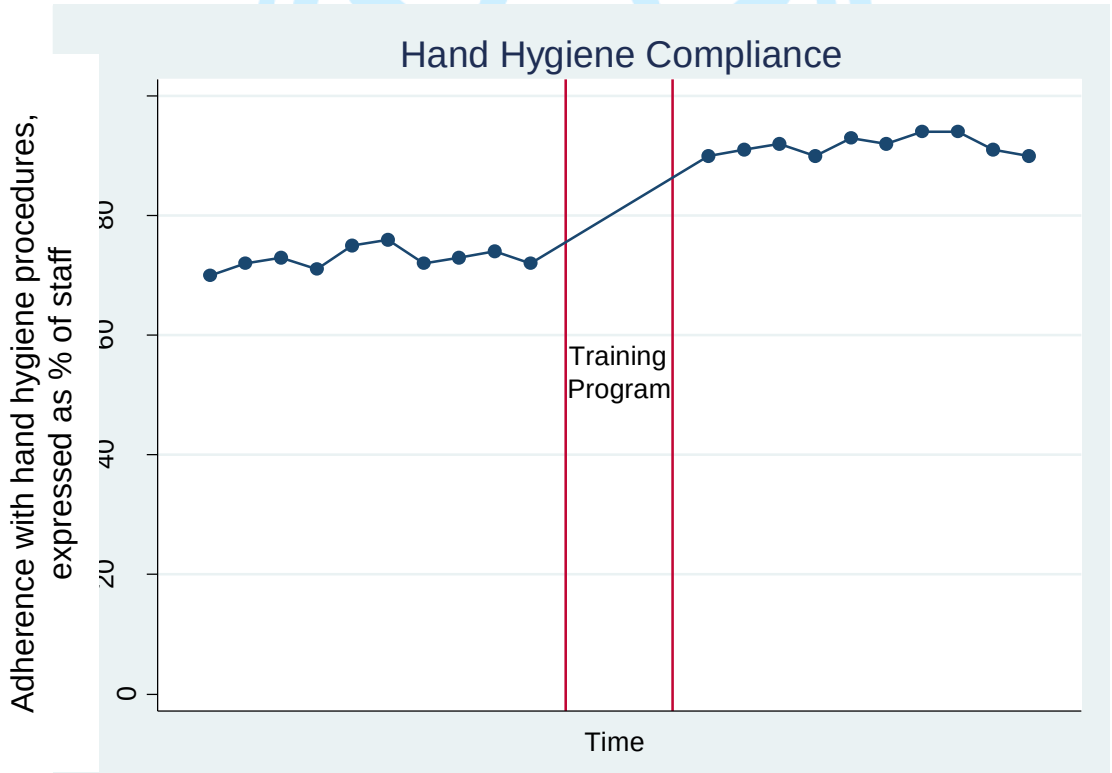


Figure 3: Adherence with hand hygiene procedures before and after a hand hygiene awareness and training program

2.6 Determine the sample size

Most often, a sample size will need to be calculated for a single proportion (see Section 2.5 for more information on choosing study design). Getting advice from a biostatistician is recommended. Remember, the sample size represents the number of completed data sets (records, participants etc) for the survey, not the number approached or records sought (see Section 3.6 for more information on maximising the response rate). The key pieces of information required to determine a sample size for a proportion are:

- An estimate of the true proportion in the source population (this requires making an educated guess, using the literature, doing a pilot study etc). For an audit this proportion will represent how often audit criteria are expected to be met.
- Degree of confidence required (commonly 0.95, or 95%).
- Precision or the amount or error you can tolerate (commonly 0.05, or 5%).
- The size of the population from which you plan to sample. (Note – not all calculators will ask for this).

Numerous free calculators exist. for example:

- G* Power (free to download).
- Statistical software such as R, Stata or SAS.
- Online calculators.

Online calculators for a single proportion include the following:

- http://www.wrha.mb.ca/extranet/eipt/files/Sampcalcaudits.xls (recommended)
- https://select-statistics.co.uk/calculators/sample-size-calculator-population-proportion/
- http://www.raosoft.com/samplesize.html
- https://www.surveysystem.com/sscalc.htm
- http://epitools.ausvet.com.au/content.php?page=CIProportion
- http://www.sample-size.net/confidence-interval-proportion/

Example of Sample Size for a proportion

To estimate the sample size needed for a hand hygiene audit, where we anticipate 85% adherence:

- Estimate of how often the audit criteria is met = 85%.
- Degree of confidence = 95%.
- Precision = 5% (which is 85% ± 5%, the confidence interval will be between 80% and 90%).
- Population size = 1,000.

To provide an estimate of 85% of health workers being adherent with hand hygiene policy, with a 95% confidence interval of 80% to 90%, a sample size of 164 is needed. This means 164 participants would need to complete provision of data.

2.7 Choose the sampling method

In order to get a 'true' estimate for an outcome of interest in a population, the entire population would need to be surveyed. This is usually not possible due to time and resource constraints. Therefore, a ***sample*** will need to be selected. The sampling method chosen must be clearly described and justified. The aim is to maximise the ability to generalise survey results to the 'whole' population of interest (the 'source' population). Getting this right is a combination of first identifying the source population, choosing a sampling method and also estimating the sample size, as the sample chosen must also be sufficiently large to enable confidence in the results (see Section 2.6 for more information on determining the sample size). Obtaining advice from a biostatistician is always advised to ensure this critical step is done well. Different sampling methods are described below.

Random sampling methods (preferred where feasible)

The process to obtain a sample is best (most likely to reflect the whole population of origin) when it is truly random, that is, every individual in the intended population has an equal chance of being selected. When conducted well, this process should generate a study sample that is truly representative of the entire 'source' population. It requires a complete list of every member of the population of interest. Random number generation is often used, usually via a computer program or software, to guide the selection of a sample from the entire population list (see Section 2.9 for more information regarding online random number generators).

Simple random sampling

Simple random sampling is the strongest sampling method, because each person has a truly equal chance of inclusion. For example, there are 500 staff in an organisation and each can be assigned a number from 1 to 500. A total sample of 150 is planned for the survey based on a sample size calculation. Therefore 150 random numbers are chosen and matched to the names.

Advantages: Simple and straight forward method offering best chance of generalisability to the wider (complete or source) population.

Disadvantages: Requires a complete 'denominator' listing of all instances in the population to be sampled.

Systematic sampling

This method selects people at regular, pre-defined intervals based on the sample size required and the total size of the population.

For example, if there are 5,000 people in the population of interest and a total sample of 200 is planned for the survey, the sampling interval would be determined by 5,000/200 = 25. Every 25th person is therefore chosen from the list of 5,000 people. The sampling begins by randomly choosing the first person. Then every 25th person is chosen until the total of 200 participants is reached.

Advantages: Simpler method for large populations compared to simple random sampling and provides a spread across the sample. This is because it is easier to select every 25th person (as an example), than to keep referring back to a list and try and pick number 7, 16, 36, 82…

Disadvantages: Requires a logically homogenous list of participants prior to commencing the sampling. Any pattern within the sample undermines the prospect of random selection. For example, selecting names from a phone book would not be suitable for this method.

Stratified sampling

This method is useful when the outcome is likely to vary between groups (or strata) and when it is important that each group has representation in the total sample. Stratified sampling first divides the population into groups (or strata) who share similar characteristics. Commonly considered strata include gender, age, education level, postcode or socioeconomic group. A sample is then chosen from within each stratum (ideally randomly).

For example, in a survey of heart disease the population is stratified by gender to ensure there is equal male and female representation. Equal numbers are taken from each stratum.

Stratified sampling is useful when groups/strata within the sample require proportionate representation similar to the wider population. This technique ensures that the proportions of the different groups (or strata) present in the population are represented as the same proportions in the study sample.

For example, a survey is planned of nursing staff from three wards. Each ward has different numbers of nursing staff. The investigators wish to have the same proportions from each ward in the study sample. Table 1 represents the number of staff and the sample number chosen.

Table 1: Worked example of recruitment targets used during stratified random sampling

Ward	**Nursing staff employed**	**Sample numbers** **(20% of the total sample)**
A	100	20
B	120	24
C	160	32
Total	**380**	**76**

Advantages: This method has proportionate representation of each stratum, thereby making it similar to the total population. It improves generalisability compared with non-stratified sampling from the total population of all strata combined.

Disadvantages: Prior knowledge of the sample characteristics and relationship to the survey outcomes is required.

Cluster sampling

Where similar units exist within a population to be sampled (e.g. schools, towns) these are called clusters. If the researcher knows that these clusters are fairly similar (e.g. inner-city public schools, or similar sized country towns in the same jurisdiction), then it is possible to (ideally) randomly select clusters for the study sample. This method aims to represent the overall population, using fewer resources compared with sampling from the entire population.

For example, a study of grade 5 students from suburban Perth public primary schools is planned. A total of 20 schools are chosen at random from the complete list of all schools fitting these criteria. Students can then be sampled within each selected cluster.

Advantages: Simple, more convenient and cheaper to administer compared to other forms of sampling when tackling large populations, whilst maintaining reasonable representativeness.

Disadvantages: Clusters must be sufficiently similar ('intra-cluster correlation coefficient' must be known from the scientific literature or estimated). Overall, less representative than sampling from the entire population of interest.

Mixed sampling methods

Often used in large epidemiologic studies of populations.

Non-random sampling methods

Non-random sampling means some people have no chance of being selected, resulting in a non-representative sample. These methods must be avoided for quantitative surveys including audits wherever feasible.

Purposive sampling

This is known as judgement or selective sampling. In this approach the researcher chooses a sample that suits their needs, such as people with certain characteristics. Although purposive sampling may be appropriate for qualitative research (e.g. inviting participants with type 2 diabetes to attend a focus group of people to determine the widest possible range of ideas for future research), it should not be used for quantitative surveys.

Convenience sampling

Convenience sampling involves choosing participants based on ease; for example availability and willingness to participate. The problem with this approach is that the people ultimately included are likely to be systematically different from those not included, due to recruitment or volunteer bias, resulting in a non-representative sample.

For example, a survey involves approaching people in a busy clinic that runs Monday to Friday, 9 am to 5 pm, to ask about their satisfaction with the service provided. The researcher is only available on Monday morning between 9 and 11 am. He avoids the mother with three children under five years of age who are all crying, the father on his phone talking in a loud, angry voice, and the family all speaking in another language. The researcher chooses to approach the relaxed looking mother with a teen on their phone.

2.8 Choose the method for survey delivery

The choice of survey delivery method impacts on the likelihood of collecting high quality data and the response rate. Whilst more resource-intensive, prospective data collection is more likely to result in a complete, high quality data set and provide the opportunity to collect the exact data that are needed. Retrospective data collection means relying on the required data having been collected previously, for example, recorded in medical records. Whilst generally cheaper and easier, this approach

almost always results in an incomplete data set, or less rich data than would be ideal.

There are many options to consider for prospective data collection and each has advantages and disadvantages including time required, cost, outreach and response rates (Table 2). Choice depends on the resources available, and the population to be surveyed.

Table 2: Summary of advantages and disadvantages for different administration processes

	Phone	Mail	Web/electronic	Face-to-face
Cost	$$	$	$	$$$
Time	+++	+	+	+++
Flexibility*	+++	+	+	+++
Response	60 to 80%	70%	?30 to 60%	80%

* ability to respond to queries participants might have about the questions

Weblinks to source information

http://www.cdc.gov/hepatitis/partners/Perinatal/PDFs/Guide%20to%20Life%20Appendix%20C.pdf

Phone

Advantages: Useful for reaching geographically diverse populations. Open ended and probing questions are possible.

Computer Assisted Telephone Interviews can also be very helpful (save time and allow participants to skip irrelevant questions). They can be combined with trained interviewer questions.

Disadvantages: Limited to those with phones (relying on landline use is definitely problematic). Difficult to engage (busy) health providers. Response rates are variable.

Complex to implement (setting numbers of call backs, allowable times, ethical issues) and requires consistent training and monitoring of data collectors to ensure data are high quality.

Mail

Advantages: Can reach anyone – great for geographically diverse areas. Good response rate (if provide stamped, self-addressed envelope). Very efficient for cost.

Disadvantages: Requires accurate addresses (highlights the importance of collecting multiple contact details). Unable to provide clarification so questionnaires need to be extremely well written, and easy to use with clear explanatory notes.

Questions can be skipped or misinterpreted resulting in incomplete or difficult to interpret data sets. Not good where literacy or language is an issue. Data entry from paper to electronic format is time consuming.

Web-based

Advantages: Popular and effective in reaching geographically diverse populations and busy providers. Cheap. Easy to use on the run if the application is mobile-enabled. Data extraction and analysis easier than working from hardcopy datasets. Simple analysis sometimes provided by services or software.

Some programs are free, easy to use, secure and offer ability to build in data checks and balances to improve data quality and security (e.g. REDCap; see: https://www.project-redcap.org/).

Disadvantages: Target audience requires internet access/computer knowledge, therefore restricting the eligible population. Inability to check against original paper records. Response rates difficult to calculate but often very low (low 'care factor' – hit delete, don't complete all questions). Data security may be uncertain (e.g. Survey Monkey data are not owned by the researcher and stored in an uncertain location - not on site, not recommended for any confidential staff or patient data, banned by many organisations - therefore always check whether program use is allowed).

Face to face investigator-administered questionnaires

Advantages: Good for low literacy or special populations. Open-ended and probing questions possible. Excellent response rates. Able to provide clarification, explanation or correction. Less missing data.

Design and layout somewhat less critical, contextual understanding and greater flexibility possible.

Disadvantages: Not great for busy health practitioners. Most expensive (resource intensive as requires staff and training of staff). Requires skilled interviewers with consistent training plus monitoring for potential interviewer bias.

Impractical if requiring frequent data collection or surveying a wide geographic area.

Self-administered questionnaires

Advantages: Relatively cheaper. Does not introduce interviewer bias. May permit a larger sample. Easier for a sample with a wide geographic distribution.

Disadvantages: Low response rate (~40% common). No chance for clarification or correction. Higher rate of missing data. Need very clear instructions. Poor literacy will compromise response.

2.9 Additional resources

Audit methodology – weblinks

A guide to using data for health care quality improvement (The Victorian Quality Council, 2008)
https://www.aci.health.nsw.gov.au/__data/assets/pdf_file/0006/273336/vqc-guide-to-using-data.pdf

About Clinical e-Audits for GPs (The Australian National Prescribing Service - how to audit pharmacological management of patients and relevant standards)
https://www.nps.org.au/about-clinical-e-audits-for-gps

Case-control and cross-sectional studies (BMJ)
http://www.bmj.com/about-bmj/resources-readers/publications/epidemiology-uninitiated/8-case-control-and-cross-sectional

Clinical audit samples (Australian and New Zealand College of Anaesthetists)
http://www.anzca.edu.au/fellows/continuing-professional-development/handbook-and-resources/clinical-audit-samples

Effect size statistics, power, and sample size calculations (The analysis factor)
http://www.theanalysisfactor.com/resources/by-topic/effect-size-statistics-power-and-sample-size-calculations/

How to: set up an audit sample and plan your data collection (University Hospitals Bristol, 2009)
http://www.uhbristol.nhs.uk/files/nhs-ubht/5%20How%20To%20Sample%20Data%20Collection%20and%20Form%20v3.pdf

Planning and conducting a survey (BMJ)
http://www.bmj.com/about-bmj/resources-readers/publications/epidemiology-uninitiated/5-planning-and-conducting-survey

Practical tools for international development (Tools4Dev)
http://www.tools4dev.org/resources/how-to-pretest-and-pilot-a-survey-questionnaire/

Standards for general practice (Royal Australian College of General Practitioners, 2017)

https://www.racgp.org.au/running-a-practice/practice-standards/standards-5th-edition/standards-for-general-practices-5th-ed

Survey methods (Queensland Government, 2017)
http://www.qgso.qld.gov.au/about-statistics/survey-methods/

Survey techniques; relative advantages and disadvantages
http://www.cdc.gov/hepatitis/partners/Perinatal/PDFs/Guide%20to%20Life%20Appendix%20C.pdf

Audit methodology - peer-reviewed papers

Boynton, P. (2004). Administering, analysing, and reporting your questionnaire. *BMJ*. 328(7452): 1372-1375.

Kaplowitz, M.D., Hadlock, T.D., Levine, R. (2004). A comparison of web and mail survey response rates. *Public Opin Q.* 68 (1): 94-101.

Kelley, K., Clark, B., Brown, V., Sitzia, J. (2003) Good practice in the conduct and reporting of survey research. *Int Journal for Quality in Health Care,* 15(3):261-6.

Nulty, D.D. (2008). The adequacy of response rates to online and paper surveys: what can be done? *Assess Eval Higher Educ.* 33 (3): 301–314.

Szolnoki, G., Hoffmann, D. (2013). Online, face-to-face and telephone surveys—Comparing different sampling methods in wine consumer research. *Wine Economics and Policy*. 2(2): 57-66.

Van den Broeck, J., Argeseanu Cunningham, S., Eeckels, R., Herbst, K. (2005). Data cleaning: detecting, diagnosing, and editing data abnormalities. *PLoS Med.* 2(10): e267.

Code for random number generation in Stata and Excel

Excel Formula	
RAND()*(b-a)+a	Generates a random number (integer) between a and b
Example	
INT(RAND()*(500-0)+0)	Generates a random number (integer) between 0 and 500.
Drag the fill handle down to fill 50 cells	50 random numbers are generated
https://www.techonthenet.com/excel/formulas/rand.php	

Stata coding	Description
generate *variable* = floor((*b*–*a*+1)*runiform() + *a*)	generates integer random numbers between a and b
Example	Generate a random selection of 50 numbers out of a total list of 500
Clear	Start with a clear dataset, by clearing all data from memory
set obs 50	Tells Stata to generate 50 numbers
generate RandomNumber= floor((500-0+0)*runiform() + 0)	Generates random numbers (integer) between 0 and 500. The variable will be called RandomNumber
list RandomNumber	Lists the 50 numbers generated
https://blog.stata.com/2012/07/18/using-statas-random-number-generators-part-1/	

Populations and samples

Populations and samples (BMJ)
http://www.bmj.com/about-bmj/resources-readers/publications/statistics-square-one/3-populations-and-samples

Random number generation - online options

https://www.random.org/integer-sets/

https://www.randomizer.org/

https://www.randomlists.com/random-numbers

Weblinks to The Research Education Program content

Data collection and management (CAHS, Australia, 2019)

Research fundamentals (CAHS, Australia, 2019)

Sample size (CAHS, Australia, 2019)

Survey design and techniques (CAHS, Australia, 2019)

https://cahs.health.wa.gov.au/Research/For-researchers/Research-Education-Program/Past-seminars

Section 3: Questionnaire design and maximising response rate

Before designing a data collection instrument or questionnaire (also sometimes called an audit tool):

- Look at the key questions and objectives and make sure all variables/responses to be collected are relevant to these.
- Check whether a similar survey has been conducted previously (literature review, peers, institutional databases etc). It may be feasible to adapt existing questions for this project, saving time and ensuring the ability to compare with previous studies.
- Know it is normal to have multiple iterations in order to get a well-functioning questionnaire. Always leave enough time for this process.

In order to maximise response rates for a questionnaire, remember:

- Keep it as short as possible whilst still answering the study objectives well.
- Have a logical flow of questions, from the least to most sensitive, and the general to the particular.
- The questionnaire must be clear, simple, easy to read and well-presented. The following basics are useful:

- Use consistent font, free from typographical errors.
- Avoid vertical (sideways) text.
- Align text, boxes, and spaces.
- Run questions straight down the page, not left to right, left to right.
- State units clearly
- Provide explanatory or clarification notes for questions.
- Provide enough space for answers to open-ended questions.

Be sure to include:

- The interviewer name or code.
- Contact and return details.
- Numbering for every question and page.
- A unique identifier on every document/page.
- Categories that are mutually exclusive and collectively exhaustive.

Also, see Section 9 (References) for details of the following relevant readings Boyton (2004), Leung (2001) and Stone (1993).

3.1 Open vs. closed questions

Why do we care about this? Whilst both types of question are useful, there are important differences in response rates, entering and analysing data, the resources required, data quality and potential biases.

Open-ended questions: These are used more often in qualitative studies on small numbers of participants. For example, 'How would you describe your health?'

The advantages of open-ended questions are that the data produced are rich and result from unprompted ideas, with no limits on responses. They can be used to generate the widest possible number of response options for quantitative studies, or to generate hypotheses for future study.

Their disadvantages include greatly increased difficulty, time and cost of summarising themes and analysing responses (data recording, entry, interpretation and analysis). Open-ended questions work less well for self-administered questionnaires, require high levels of inter-interviewer consistency to avoid introduction of bias, and can result in responses that do not address the question as intended.

Closed questions: Closed questions have pre-determined, limited option responses. For example, what is your gender? (1 = male, 2 = female, 3 = other - specify, 4 = prefer not to say).

The advantages of closed questions include that they:

- Are quick and easy to answer.
- Force categorisation, so data are easy to enter and analyse and enable comparisons.
- Clarify the meaning of the question by providing choices.
- Minimise intra-interviewer bias.
- Maximise inter-interviewer consistency.

The disadvantages of closed questions include that they:

- Can't include all possible responses.
- Require prior knowledge of likely response options.
- Don't necessarily capture respondents' thoughts.
- Produce data that are less 'rich' and less likely to generate new ideas.
- May produce random responses, particularly where there is a 'middle' option, or the person does not understand the question.

Suggestions to avoid common problems with wording of questions

- Use clear, simple “plain” language. Use vocabulary that a 12 year-old would understand.
- Avoid using medical jargon, abbreviations, or acronyms.
- Avoid demanding or sensitive questions.
- Ensure questions are not ambiguous. For example:
- Avoid words like ‘often’, ‘regularly’ ‘usually’ – they mean different thing to different people.
- Stipulate time frames (e.g. hours, days, weeks, months, years).
- Specify the level of accuracy/units you want (e.g. kilograms to one decimal point).
- Ensure questions are grammatically correct.
- Make sure questions are objective. For example:
- Avoid labelling and value judgments.
- Avoid leading questions.
- Don’t ask two questions in one.
- Avoid double negatives to minimise confusion and incorrect responses.
- Don’t over-estimate memory, or useless data will result.

3.2 Writing response options

When writing closed questions, response options must be:

- Mutually exclusive (non-overlapping).
- Collectively exhaustive (i.e. cover the full range of possibilities), for example yes, no, don’t know, not applicable

The following example demonstrates a number of problems:

Question: How many times do you eat with your children per week?

[ ] 0 to 1

[ ] 1 to 3

[ ] 3 to 5

[ ] 5 to 7

First, categories overlap, making analysis impossible. Second, the units are not defined (?meals at the table, ?snacks in the car), and there could be other options for a response like ‘don’t know’, ‘not applicable’, or more number categories, depending what the unit of response are. Overall, the data from this question would need to be removed from the analysis, due to the responses being uninterpretable.

Remember:

- Include an ‘**other (specify...)**’ or ‘**don’t know**’ response option. There is nothing more annoying for a respondent than being forced into an incorrect response because the option they need is not provided.
- A neutral category (e.g. middle option in a 5-point Likert scale) may mean:

- No strong feelings.
- Ignorance or uncertainty.
- Questionnaire fatigue and just wanting to get it finished.

Do not use ‘tick all that apply’

This approach cannot distinguish between missed questions and those that are deliberately not chosen. For example, when asking the following question:

Which of the following increases the chance of a heart attack?

Rather than asking the respondent to ‘tick all that apply’

- Smoking
- Being overweight
- Stress

It would be preferable to offer the following response options (please circle):

Smoking	Yes	No	Don’t know
Being overweight	Yes	No	Don’t know
Stress	Yes	No	Don’t know

3.3 Skips

Skips improve ease of questionnaire administration/completion - as long as they are correct!

E.g. Q1. Did you attend the questionnaire session? [ ] Yes [ ] No →Q8

3.4 Minimising data errors and bias

Remember that the aim of an audit or survey is to obtain data with maximum validity and reliability. This can be achieved through good planning (sampling strategy, analysis planning and questionnaire design) but also by setting up a robust database, training project staff and piloting all processes before “real” data are collected. Once data are collected, data must be cleaned to ensure they are as complete and correct as possible prior to analysis. Without cleaning, data are likely to be poor quality or even uninterpretable, resulting in wasting of respondent, interviewer and analyst time, as well as all the other resources required for a project to reach completion.

In general, the following are factors need to be considered when undertaking a survey in order to minimise data errors and bias:

- Development of the sampling plan.
- Questionnaire design.
- Piloting the questionnaire.
- Database design.
- Staff training.
- Data collection and monitoring.
- Data cleaning.

Instructions

Instructions are useful for those administering or completing questionnaires, as they reduce misinterpretation and improve consistency of administration and completion of questions. For example, if asking: Do you have a history of cardiovascular disease? Consider giving the following examples: (angina, heart attack, stroke……)

3.5 Pilot the data collection instrument and delivery method, and train study personnel

Piloting all aspects of the survey should be considered essential and only missed at the researcher’s peril! It is a critical step in detecting mistakes and will give the best possible chance of successful survey completion. It is also useful for ensuring each variable has all possible options (collectively exhaustive) or categories available; see Section 3.2 for more information on writing questions and response options).

When running a short pilot study, test the whole process as well as the data collection instrument. Test the questionnaire on people or a few records similar to those that will be in the actual survey, but who will not be included. Also ensure it is

tested on someone who knows absolutely nothing about the topic in question, as well as someone who is good at providing critical advice on layout and design.

Piloting a questionnaire will involve checking:

- Does the questionnaire flow logically?
- Is it verbose and difficult to read?
- Do the questions mean the same thing to different people?
- Has something important been left out?
- Does the questionnaire make sense?
- Can data be recorded accurately?
- Do categorical variables have mutually exclusive and collectively exhaustive categories? Do new categories need to be created? Are units and time frames specified and consistent?
- Do the skips work?
- Is the layout efficient and easy to read?
- Can I enter the data easily?
- Does the database function as expected?

Training and monitoring

Training and monitoring staff who collect, store and enter data is paramount for maintaining data quality. This includes regular checking of interview technique if conducted face-to-face or via telephone. Writing Standard Operating Procedures for making contact with participants, administering questionnaires, cleaning, entering, storing and securing data, and use of metadata will ensure consistency between staff. Monitoring actual practice improves consistency over time. Standard Operating Procedures should also include the process for seeking advice when uncertainty exists, and how to communicate this (including updating of the Standard Operating Procedure) when resolved. See Section 3.7 for more information on issues related to data storage.

Interviewing tips for minimising bias

- Use the same approach every time.
- Always ask questions in the same order.
- Additional information must be the same for all.
- Never 'paraphrase.'
- Never leave out information.
- Never make assumptions.
- Note formal skips and prompts.
- Never leave blanks – have a code for missing data.
- Don't 'complete' partial answers e.g. day and month only for date of birth.
- Seek clarification where needed.
- Have clear instructions for corrections.

There are a number of types of bias that are particularly pertinent to survey administration. It is important to be able to report how these were minimised or estimated when discussing potential limitations of the study results.

Interviewer or observer bias

Inter-observer bias: More interviewers mean more variation in how questionnaires are delivered.

Intra-observer bias: 'Spin', or variation from wording, or deviation from the standard order of a questionnaire can occur over time.

A well written questionnaire that is piloted following training of project staff can minimise inter- and intra-observer bias, particularly if monitoring of staff occurs through the life of the project.

The Hawthorne effect

This occurs where participants change their behaviour due to their awareness of being studied. For example, participants are asked to report on their handwashing practices, and simply being in the study makes them want to report better or 'expected' practices.

Recall bias

This occurs where participants have differential memory of an event. For example, participants who have recently had pneumonia are more likely to remember whether they have had their annual influenza vaccine than those who have not had pneumonia.

3.6 Maximise the response rate

What is the response rate?

The response rate is the number of people who completed the survey, expressed as a proportion of the total number of eligible people approached. This final number of ***completed*** surveys should be equal to or greater than the calculated sample size (see Section 2.6 for more information on determining sample size). Hence when estimating the number of people to be approached for a survey, it is important to allow for those who do not meet eligibility criteria, those who decline to participate, those who are lost to follow up or who do not complete the survey and so on. When reporting the response rate, it is useful to outline the number of people (see Section 4.1 for more information on reporting response rates):

- Approached.
- Eligible (versus 'out of scope').
- Who agreed to participate.
- Who completed the survey.

It can sometimes be difficult to determine how to categorise these outcomes depending on the survey method chosen. In general, the following responses are treated as 'out of scope':

- Mail-out surveys that are returned as 'not at this address'.
- Uncontactable telephone numbers.
- Unoccupied dwellings for face-to-face surveys

When planning a survey, it is useful to pre-define responses that will be considered 'out of scope' so that a consistent approach is taken, and corresponding records are kept in the dataset.

Why is a high response rate important?

In order to achieve as representative sample as possible, a survey must have a strong sampling method (see Section 2.7 for more information on sampling methods) ***and*** a high response rate. Whilst there is no 'magic' proportion to be achieved, a response rate of at least 80% is desirable. The reason for this is that the non-responders may be systematically different from the responders – and if included, could change the end result. This is called 'response bias'.

For example, if only 50% of eligible participants completed a survey about whether they liked a clinical service provided and all of them liked it, you might assume that 100% of respondents liked the service. However, if the other 50% all hated the service, if included, overall only 50% actually liked the service; a very different result. It is always possible to undertake a 'sensitivity analysis' such as this following completion of a survey, to see how different the result might be under such extreme conditions when the missing respondents are included, therefore providing a range of possible outcomes. Obviously the lower the response rate, the less confidence

there can be in the data reflecting the true outcome for the source population. The more different the 'missing respondents' are likely to be, the higher the required response rate.

A good example is referenced by the following article: Data Analysis Australia (see weblink to source information below) taken from the National Indigenous Preschool Census conducted by Data Analysis Australia for the Australian Government. 'The aim of this census is to measure the participation of Aboriginal and Torres Strait Islanders in pre-schooling through a survey of preschools. The preschools that are least likely to respond to the surveys are the remote, hard to contact, community-run ones - precisely those with the highest numbers of Indigenous children. Consequently, a response rate less than 95% is considered poor!'

Weblink to source information

Data Analysis Australia (2005)
http://www.daa.com.au/analytical-ideas/response-rates/

How to maximise response rates

Note, the resources required for the following approaches for improving response rate need to be weighed against the increase that is likely to be achieved.

Pre-contact: Phoning or sending a letter, email, text or fax to inform respondents about the survey or to set up an interview time often helps, especially if this information comes from someone who can be trusted. Typically the organisation commissioning the survey would do this. It is important to consider the burden on potential participants. Ethics, governance or safety and quality officers can provide guidance as to whether or not the proposed approach is too onerous or intrusive.

Minimise effort for participants: A well-designed questionnaire that is hopefully of interest to participants and as short as possible is very important. Minimise the length and difficulty of the survey, simplify the format, ensure questions are not ambiguous, have clear instructions and only ask what is necessary (see Section 3.1 for more information on questionnaire design).

Follow-up/reminder plan: Phone or send a letter, email, text or fax to people who haven't responded within a certain time frame. Multiple follow-ups may be needed. It is important to ***pre-determine*** how many repeat attempts to contact eligible participants will occur, and at what times. This approach should be approved in advance and must be a trade-off between getting a high response rate, costs, and unduly harassing potential participants. Unique identifiers on each survey form are generally required, so reminders can be sent to those who haven't responded, rather than a general notice to all people in the sample (this is obviously annoying for those who have already responded). Unfortunately, identifiers on surveys can lead to an initial lower response rate as some people may interpret this as jeopardising the confidentiality of the survey.

Incentives: Some offer a prize to people who respond to the survey or an up-front gift/payment/voucher to everyone approached. The form of the incentive may cause bias because particular groups may find it more appealing than others. Incentives that require people to identify themselves on the forms can possibly lead to untruthful answers or a lower response rate. Wherever incentives are being considered, they must be chosen carefully and approved in advance after consideration of ethical issues.

Timing: Think carefully about optimal timing. For example, conducting a survey of school-aged children during school holidays would very likely to fail. Timing of telephone interviews is very important (day of week, time of day). A pilot study can identify suitable times, or optimal times can be clarified at the initial contact.

Tips for mail-out surveys:

- Provide a cover letter that outlines the purpose of the survey and provides contact details.
- Specify a return date, which is relatively short (e.g. 2 weeks) to minimise the chance of a survey being forgotten).
- Provide a self-addressed envelope (pre-paid if possible).
- Ensure a follow-up plan is in place.
- Ensure individuals can be linked to their data with a unique identifier (see Section 4.3 for more information on setting up the database).
- Get other contact details.

Weblink to source information

Data Analysis Australia (2005)
http://www.daa.com.au/analytical-ideas/response-rates/

3.7 Additional resources

Weblinks to The Research Education Program content

Data collection and management (CAHS, Australia, 2019)

Research fundamentals (CAHS, Australia, 2019)

Survey design and techniques (CAHS, Australia, 2019)

https://cahs.health.wa.gov.au/Research/For-researchers/Research-Education-Program/Past-seminars

Section 4: Data and database considerations

4.1 Choose the variables for data collection and plan the analyses

A survey must be as succinct as possible whilst addressing all the pre-defined objectives. Planning the analysis before designing the data collection instrument (e.g. paper or electronic form) ensures that the questions included capture all the necessary variables to answer them, but no more. Collecting data without a clear and pre-specified purpose reduces efficiency, including investigator and participant time, and time required for data entry and cleaning. Having a very clear question, underpinned by a thorough literature review, and discussion with peers and other stakeholders in the field, can ensure all necessary information is collected, and comparisons can be made with other similar studies where appropriate.

Ensure relevant definitions of outcome measurements and for case selection are clear in terms of person (or clinical factors), place and time.

For example: Proof of staff influenza vaccination via the Australian Immunisation Register or sighted personal record card, with an approved/licensed influenza vaccine, given between and including 1 March to 31 October 2019.

Weblink to source information

Norms and Standards in Epidemiology: case definitions (Pan American Health Organisation, 1999)
http://www1.paho.org/english/sha/be991norms.htm

Overview of common variables and analysis

Variable lists generally include (depending on the type of survey not all may be applicable):

- Unique identifiers for each participant or record (essential).
- Name, address, contact details, date of birth (store separately, link to unique identifier).
- Participant characteristics (e.g. age, gender, comorbidity etc).
- Participant inclusion and exclusion criteria.
- The recruitment process (e.g. time points, location, number approached).
- Time variables such as; enrolment, visits, exit, loss to follow-up.

- Consent variables.
- Variables for who is collecting the data
- Important subgroups.
- Primary and secondary hypothesis variables.
- Other outcomes of interest.
- Potential confounding and exposure factors (see Section 4.6 for more information on confounding)

An analysis from a survey would generally include:

- Simple descriptive data.
    - Participation numbers: number approached, eligible, out of scope, refused, consented, dropped-out, lost to follow-up, completed. For a simple audit, this might be represented by number of hospital records requested, number available, number completed.
    - Overall numbers for outcomes of interest, means, median, proportions.
    - Participant characteristics (e.g. gender, age, comorbidities etc). or characteristics of a clinical condition (e.g. severity, duration etc) or characteristics of an aspect of practice (e.g. completeness, location, staff age, experience, training background).
- It may also be appropriate to include:
    - Simple tables (consider use of t-tests, chi-squared test etc). Examine variation in the outcome of interest between different groups.
    - Adjustment for potential confounders using multivariable analysis (see Section 5 for more information on analysis).

If unsure where to start, consider:

- Writing the title(s) of the intended reports or peer-reviewed papers.
- Draft the abstracts for the reports/papers.
- Outlining tables or graphs needed (‘dummy tables’).
- Look at other studies/speak with content experts and stakeholders.
    - Validated data collection tools already that may be adapted or used with approval?
    - Other variables to enable comparison with existing studies?

Data types and why these matter

The variable type chosen has implications for the richness and completeness of the final dataset, as well as the resources required to enter, analyse and interpret the data. Choice of the wrong data type can severely impair the ability to analyse the data. For these reasons, always discuss the intended data set prior to commencing a project, for example with a statistician, epidemiologist, data manager or someone with data and/or content expertise.

Data are generally considered in numeric (continuous or discrete), 'string' or date formats.

Data can also be considered as categorical (nominal), ordinal and discrete.

Numeric data

Numeric variables should be used for any data for which mathematical calculations are planned. Numeric codes are best for many categorical variables, as they make data entry easier, plus many statistical procedures will only work with data stored in numeric format. There are different types of numeric variables: continuous versus discrete and categorical (or nominal), ordinal and interval.

Continuous variables: Continuous variables have an infinite number of possible values within a selected range e.g. temperature, height or weight. Data can be ordered along a scale or within a range, and the distance between each point is equally spaced. It is important to choose an appropriate unit and number of decimal points, and to remain consistent with this for similar variables (e.g. birth weight, weight at time of study) to reduce errors and enable calculations to be made.

Discrete variables: Discrete variables take specific numeric values and are based on counts. For example:

- An Apgar score must be a whole integer between 0 and 10.
- The number of students in a class must be a whole number.

Nominal variables: Nominal variables have categories with no inherent ordering. They are binary if only two options (e.g. yes or no). Frequently codes are still used (e.g. yes = 1, no = 0) to enable easier data entry, reduce errors, and facilitate analysis. For example:

Marital status: 1 = single, 2 = married, 3 = defacto, 4 = separated, 5 = divorced, 6 = widowed, 7 = other-specify, 8 = prefer not to say.

Gender: 1 = male, 2 = female, 3 = other - specify, 4 = prefer not to say.

Ordinal variables: Ordinal variables have categories on an ordered scale, conveyed by assigned codes to assist data entry, reduce errors and facilitate statistical calculations. The different categories are not necessarily the same size or distance apart. For example:

Rating of an experience: 1 = poor, 2 = fair, 3 = good, 4 = very good, 5 = excellent.

Socioeconomic status: 1 = low, 2 = medium, 3 = high (criteria provided).

Incapacity rating: 1 = mild, 2 = moderate, 3 = severe.

Education level completed: 1 = primary school, 2 = secondary school, 3 = etc

Interval variables: Interval variables are ordinal variables with categories of equal size. Often continuous variables are recalculated into interval variables. For example, income level may be broken up into equal sized, non-overlapping categories

Why does it matter if a variable is nominal, ordinal or interval? 'Statistical computations and analyses assume that the variables have specific levels of measurement. For example, to calculate an average, a variable must be an interval variable. It would not make sense to compute an average of a nominal variable like gender or eye colour because there is no intrinsic ordering of the levels of the categories. The average of an ordinal variable like education level would also not make sense because the spacing between the different levels is uneven.'

The analyses of 'in between' variables, such as those on a 5-point Likert scale (e.g. satisfaction surveys: 1 = strongly disagree, 2 = disagree, 3 = neutral, 4 = agree, 5 = strongly agree), are more difficult. If confident that these are equally spaced, in theory an average could be applied.

Weblink to source information

What is the difference between categorical, ordinal and interval variables? (Institute for Digital Research and Education, University of California, Los Angeles, USA) https://stats.idre.ucla.edu/other/mult-pkg/whatstat/what-is-the-difference-between-categorical-ordinal-and-interval-variables/

String variables

String variables are 'free entry' variables that can contain text and numbers. It is ***not*** possible to perform mathematical or statistical operations on string variables. They should therefore not be used if a numeric variable can be used instead. Most commonly, string variables are used for names and addresses only, or in qualitative research, for data that is too complicated to categorise, such as long comments. Many programs have limitations on the length of string variables. For simple text outcomes, it is always better to assign codes from the outset to enable calculations and reduce errors. For example: Yes = 1, No = 0

Some string variables can be converted to a numeric code for analysis later, but this requires complete consistency with spelling, case etc. For example: Yes could be written as Y, Yes, yes, or misspelled.

In general, string variables should be avoided wherever possible by good early planning of variable types.

Date variables

When using dates, in order to enable calculations of time intervals, it is very important to use a consistent date format in the data entry package. Never enter dates as string variables or as separate day, month and year variables. Use of a date format ensures only valid dates will be entered, particularly if acceptable ranges are set around them.

Date formats vary around the world. Both day-month-year and year-month-day are logical date formats that are easily understood. Choose one and use it consistently throughout the database.

Raw data are important

It is preferable to collect and enter data in a 'raw' format and categorise later. This allows for checking against original records and reduces human error. For example, collect and enter variables 'participant date of birth' and 'date of review' in a consistent format date variable. Later use an analysis program to calculate new variables 'age' and 'age groups'.

Another example would be to collect three measurements for infant head circumference with the intention of using their average in further analysis. It is best to keep the three individual measurements, and later use an analysis program to calculate the average.

The unique identifier variable

Every participant or case must have a unique identifier as a variable in the database. This field is usually the very first variable in every database. It should be present on every page of every data entry form related to an individual or case.

Unique identifiers are used to refer to specific records/people in the database without using identifying information such as names and addresses. It is very important that these identifying variables are kept separately from the main database and are able to be linked to individuals using the unique identifier. Linking should only be performed by key people with the clearance to do so. Unique identifiers are also used to link data between tables in a database.

The unique identifier for a participant is usually a sequential number assigned by the researcher. A hospital record number or date or birth should never be used as the unique identifier as these can easily identify the participant and their personal details, with potential for breaches in confidentiality.

4.2 Data dictionary

A data dictionary is an essential list of all the variables in the database, their characteristics and relationships to other variables. It underpins the development of the database. It is often written as a simple word document. It usually includes:

- Table names (e.g. data collected on visit 1).
- Variable names.
- Variable descriptions.
    - Meaning, data type, units of measure.
    - Validation/coding rules.
    - Codes for categorical variables.
    - Ranges for continuous data and dates.
    - Codes for missing and queried data.
- Relationships between tables.

Examples of variables as data dictionary entries

- Studyid.
    - Essential to have a unique identifier for each record.
    - Identification numbers should be written on every form.
    - Specify the range of Identification numbers on the coding sheet.
    - There should be no missing values.

This variable would appear as:

Variable name	Description	Data type	Values/rules
Studyid	Participant's unique study ID number	Number	Must be unique 100 – 200 allowed

- Weight at 2 year visit (weight2)

- Continuous variable.

- Include unit of measure (kg).

- Specify plausible range.

- For missing and query codes, use numbers that are outside the possible range of values for this variable.

Variable name	Description	Data type	Values/rules
weight2	Weight at 2 year visit (kg) One decimal point	Number	30 - 150 777 = query 999 = missing

Response codes

In general, response codes are useful to increase speed and accuracy of data entry, and sometimes these represent options for analysis. Codes are usually applied to each response category. For example: Yes = 1, No = 0 or Male = 1, Female = 0.

Coding information is always included in data dictionaries, which describe all characteristics of each variable and assist database use by the project team and future users.

Query codes

These are a 'code' to indicate a 'data query' encountered during data entry. A query code indicates a piece of data that needs to be checked or followed up. Data may be illegible, unlikely, unclear or inconsistent with other responses. For example, the date of birth is after date of visit, or a respondent indicates they are male but answers 'yes' to 'Have you been pregnant?'

For numeric variables, missing data is conventionally represented by 7s (e.g. 77 or 7/7/7777 for dates.

A value is chosen that is not a valid response for the field to indicate a query. For example: if someone's weight could be 77, then '777', and not '77', should be used as a query code. Alternatively, to maintain consistency, it can be preferable to nominate a large number, say 7777, as the missing code for all numeric variables in the data collection instrument.

Unique identifier fields and eligibility criteria variables must have a valid value and therefore should not have a query code. Assuming best practice of setting up a database in a true data entry package (not Excel, not Survey Monkey etc), these variables are ideally set up to be compulsory, and the 'query code' should be an invalid response.

All queries should be dealt with before beginning data analysis.

Missing data codes

Every database must have a code for missing data as an outcome for every variable. Some fields, such as unique identifiers and eligibility criteria, should never be missing. As above, assuming best practice of setting up a database in a true data entry package (not Excel, not Survey Monkey etc), these variables are ideally set up to be compulsory, and the 'missing code' should not be a valid response.

Fields with missing data can be collated later on and efforts made to retrieve missing data wherever possible. All fields should have data in them unless they are skipped because of a conditional jump. For example, if a participant says no to 'Have you been diagnosed as diabetic?', then they should not have data for the next question 'Date of diabetes diagnosis.'

For numeric variables, missing data is conventionally represented by '9' for one-digit variables, '99' for two-digit fields, and so on. For dates, 9/9/9999 can often be used as the missing code.

As for data query codes, the missing code must never be a valid response for that variable.

Missing codes are always changed to the missing value for the analysis package before analysis is commenced. For example, in Stata, this is "." Otherwise, calculations will inadvertently include these as numeric responses.

4.3 Set up the database

Always use a suitable data entry package, such as REDCap, EpiData or Qualtrics. Excel does not have the desirable features listed for data entry packages below. Whilst Survey Monkey has a few of these features, it should never be used for sensitive or identifiable data and is banned by many institutions for reasons of data security and ownership (see Section 2.8 for more information on methods of survey delivery).

Get advice from a data manager if possible regarding the set up and testing of the database. One of the key features to determine is whether the database will be in a wide or long format.

Below are desirable features of a data entry package. Packages should have capacities related to:

Data set-up

- Specify a range for continuous and date variables to reduce the chance of data entry error. This prevents improbable results being included (e.g. ranges for dates, height or weight ranges).
- Specify legal values for a categorical variable (e.g. No, Yes, Don't know, Not Applicable to be entered as 0, 1, 7, 9, respectively).
- Make data entry compulsory for a variable, so that these can never be left blank
  (e.g. for unique identifier and eligibility criteria variables).
- Specify uniqueness for some variables (e.g. the unique identifier). This ensures that no two responses can be the same.
- Build in 'skips' to ensure participants do not complete variables for which they are ineligible (e.g. males should not answer questions about pregnancy, so particular values result in a jump to a question further down the questionnaire).
- Allow more than one person to enter data simultaneously.
- Have relationships between tables (i.e. relational databases).

Data validation and cleaning

- Undertake more complex checking or consistency checks, run histograms, ranges, check reasonable date intervals, ages etc.
- Allow double data entry for a proportion of the data set and provide 'error rates'.

Data analysis

- Undertake simple analyses and produce reports.
- Be easily imported into a data analysis software package if needed.

Data security

- Allow rules to be set around who can access what parts of the data set.
- Have data stored on-site rather than in a cloud if web-based.
- Have high security.

Examples of suitable free data entry packages include: REDCap and EpiData. A paid alternative is Qualtrics. As different institutions may provide free access to suitable programs, it is useful to explore what options are available via research support or safety and quality/clinical governance staff.

4.4 Collect and enter the data

Providing all the above steps have been included, data collection should now go smoothly. However, it is still possible to detect problems with either the data collection instrument or database. If problems are evident in the data collection instrument or database, it is always best to revise and resubmit for approval as soon as possible. However, if not feasible, deficiencies must be noted when final data are interpreted and reported. In some cases, this will mean excluding certain variables. Keep records of any problems encountered during data collection that enable accurate interpretation and reporting.

Data entry

Ideally data entry occurs throughout the data collection period to ensure any residual problems with the database not detected during the pilot phase can be dealt with as soon as possible.

Double entry

Some data entry programs allow double entry of a proportion of the data (e.g. 5%, ideally by different operators), to ensure interpretation of hardcopy data is consistent.

Hard copy checks

As a further check, larger studies may randomly select 10% of the data records and have two people compare a print-out of the selected electronic data with the paper collection forms.

4.5 Clean the data

All data must be 'cleaned' before analysis commences to maximise completeness and accuracy of the dataset. Checking and cleaning data takes time so plan for this from the outset. Once the database has been 'closed', as decided by the main investigator, any data corrections must be made in a statistical package. A data manager has data cleaning skills and may be able to assist with this process.

The cleaning process must be documented and reproducible. Do not do any data cleaning 'interactively.' Instead, ideally the original dataset is imported into a data analysis package (e.g. Stata, SPSS) or is already in a package that can undertake data cleaning (e.g. REDCap) and kept intact. Cleaning processes are kept as a permanent record by writing a 'command' file that can run on the original database and save the 'cleaned' file with a new name (e.g. a do-file in Stata, a syntax file in SPSS etc). This ensures the original database is protected, and a record is kept of all changes made and why.

Do not overwrite original data or variables. It is often necessary to re-code or modify original data. It is good practice to assign new variable names to modified data and keep the original variables unchanged. The exception to this recommendation is replacing missing codes with missing values.

General steps for data cleaning

All of the following are feasible in purpose-designed data analysis packages (i.e. not Excel):

- Collate 'missing' fields and collect all possible data points.
- Collate 'query' fields and ensure correct data points are entered. There should be no query codes left before analysis – if necessary, data points are recoded to 'missing.'
- Identify implausible values.
- Undertake inconsistency or 'logic' checks – run queries on the data to pull out records that don't match certain rules defined by the project e.g. age criteria.
- Finally, replace missing codes with missing values for the analysis program (see comment * below).

Missing data

- Check for blanks. It is advisable that codes for missing data are created prior to data entry. Therefore when data are entered, the only fields that should be left blank are those with a skip, e.g. If 'No' skip next question. This makes it easy to spot fields that have been skipped in the data entry process.
- If an error is found, ideally the value should be changed to the correct value. However, if there is no record of what the value should be, the missing value code should be used, e.g. 9 = missing.
- Tell the package which values indicate missing data. Usually this means converting the numeric missing code to the program's official 'missing value'. E.g. Stata = "." (see Section 3 for more information on designing the data collection instrument).

Logic checks

- Check for duplicate records. Each record (e.g. participant) in a file must have a unique identifier. This step is unnecessary when using databases that can enforce uniqueness of a variable.
- Check for consistency between variables. Data values can depend on the value of another variable. For example, in a study of survival after a kidney transplant, information on the number of previous pregnancies is relevant only for women, who should all have a non-missing value, whilst men should all have a missing value.
- Check that all participants meet eligibility criteria.
- Check repeat measures are within a reasonable range of each other, e.g. repeated newborn head circumference measures should have small discrepancies.

Categorical data checks

Because categorical variables have pre-determined plausible values, it is simple to check if this is true.

- Produce frequency tables showing all the recorded values for each categorical variable, ***or***
- If the package allows, include statements that make explicit checks on values. For example, if gender has been coded as: 0 = female, 1 = male, 9 = missing, then the statement would assert that gender contains only the values 0, 1 or 9. If the statement fails, then the variable contains at least one dubious value. This can be located using the unique identifier.

Check continuous data

Specify lower and upper limits on what is reasonable (needed where not set up already in the database development phase). Values that fall outside this range may not necessarily be wrong. All suspicious values should be checked and any errors corrected. If a value is impossible rather than just unlikely, it should be recorded as 'missing'.

- Produce summaries showing the mean, median, variance and minimum and maximum values for each continuous variable.
- As with categorical variables, include statements to check for values below the expected minimum and above the expected maximum values.
- Produce a dot plot to easily spot any possible errors.

Checking dates

- Check all dates are within a reasonable time span. For example, in a study of year 7 students, the date of birth should be about 12 years prior to the survey date.
- Check dates are in the right order. For example, date of birth < date of 1st visit < date of 2nd visit.
- Ages and time intervals can be calculated via a statistical package using the relevant dates. Check that ages and time intervals lie within the expected range, for example, negative ages indicate data error.

Longitudinal studies

- Where the same variable is measured at several time points for each participant, it is valuable to plot each participant's sequence of recorded values to ensure that they behave reasonably.
- Check that variables that shouldn't change over time are consistent.
- Only identification numbers/participants with a baseline record should have data at later time points.

4.6 Additional resources

Confounding

Confounding variable: simple definition and example (Statistics How To)
https://www.statisticshowto.datasciencecentral.com/experimental-design/confounding-variable/

Weblinks to The Research Education Program content

Data collection and management (CAHS, Australia, 2019)

Research fundamentals (CAHS, Australia, 2019)

Survey design and techniques (CAHS, Australia, 2019)

https://cahs.health.wa.gov.au/Research/For-researchers/Research-Education-Program/Past-seminars

Section 5: Analysis

This section provides an overview of the analyses of quantitative data collected via audit or survey.

Many surveys only require simple data description and analysis techniques. Statistical methods such as calculation of confidence intervals and hypothesis testing are commonly used. Subgroup analysis can explore the data in greater detail and generate new hypotheses. Although more sophisticated statistical methods can be used it is important to seek advice to determine if this is necessary or appropriate.

5.1 Descriptive statistics

Measures of central tendency and dispersion (for continuous and ordinal data)

The choice of which statistic to report for a variable depends on its data type and data distribution (Table 3). The distribution of continuous data can be checked using a frequency histogram and will influence how to report the data and what statistical test can be used. Data may be normally distributed (Figure 4), skewed (Figure 5) or contain outliers.

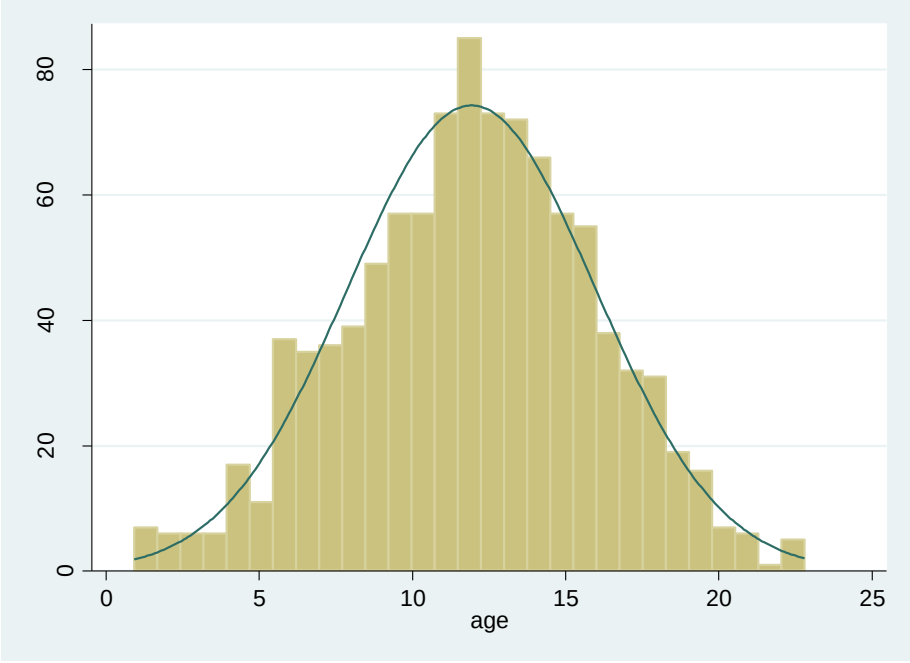


Figure 4: Frequency histogram for age (x axis) against frequency (y axis); data follow a normal distribution

Mean

Often referred to as the average, the mean is the sum of the value of each observation in the dataset divided by the number of observations.

If the distribution of the data is skewed or influenced by outliers the median may be a better choice.

Standard deviation

The standard deviation is a measure of spread of how spread out the data are for a variable. It is best used for continuous data, and when the data are not significantly skewed (e.g. by outliers). It is typically reported in conjunction with the mean.

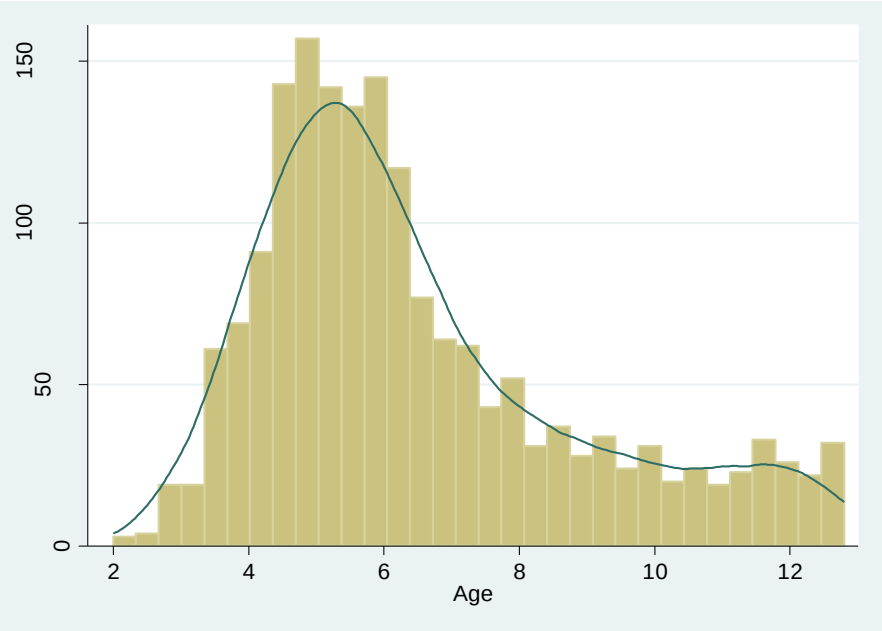


Figure 5: Frequency histogram for age (x axis) against frequency (y axis). Data do not follow a normal distribution (skewed to the right)

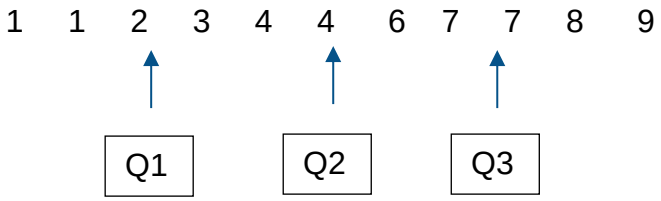
Median

The median is the number that sits halfway through the dataset. It sits exactly in the middle with 50% of values above it and 50% of values below it. First arrange the numbers in order from lowest to highest then locate the middle number. It is commonly used for ordinal data or skewed continuous data.

Interquartile range

The interquartile range (IQR) is a measure that describes the spread of the data. This is a measure of where the middle 50% of the data are located. First arrange the values in order from lowest to highest then break the data into quarters (four equal parts). Quartile one (Q1) is the first cut and quartile 3 (Q3) is the third cut. Quartile 2 is the middle cut and represents the median. It is common to report the IQR as the Q1 and Q3 (i.e. IQR 2 to 7), or by subtracting Q1 from Q3 (i.e. Q3-Q1 = 7-2 = 5). The

IQR is commonly used with ordinal data or skewed continuous data. It is reported in conjunction with the median.



Mode

The mode is the value that occurs most frequently.

Weblink to source information

Measures of Central Tendency (Australian Bureau of Statistics)
https://www.abs.gov.au/websitedbs/a3121120.nsf/home/statistical+language+-+measures+of+central+tendency

Proportions

For an audit that compares data collected against a set of criteria or standard, outcomes are most commonly reported as a proportion or percentage of adherence. It is useful to show both the percentage as well as the number for each outcome. For example, adherent: n = 40/50 (80%). It is important not to use multiple decimal points for percentages where audit numbers are small. For example, for a sample of 100, no decimal points would be reasonable.

Table 3: Summary of descriptive statistics grouped according to the type and distribution of the data

Data type	Statistic to report	Report statistic with
Continuous - normally distributed - reasonable number of observations (30+)	Mean	Standard deviation
Continuous - skewed	Median	Interquartile range and/or range
Ordinal	Median Proportion / percent in each category	Interquartile range and/or range Frequency (number)
Nominal	Mode Proportion / percent in each category	 Frequency (number)

Confidence intervals

When collecting data most projects do not collect data from the whole population. Instead a sample is taken and used to estimate a value. This estimate will not be completely accurate, and there will be some error involved. A larger sample is likely to be more accurate and produce an estimate close to the true value in the population, whereas a smaller population is likely to be less accurate.

Confidence intervals (CI) are a way of capturing this accuracy. They provide a range of values that is likely to contain the true value. The 95% confidence interval is most often used. This interval:

- Contains the true (population) value 95% of the time.
- Fails to contain the true value 5% of the time.

For example, if adherence with hand hygiene protocols was determined to be 83% with a 95% CI of 82% to 84%, this means that:

- There is a 95% chance that the confidence interval of 82% to 84% contains the true population percentage, or
- We are 95% confident that the interval 82% to 84% contains the true population percentage

Having a narrow CI is ideal as it means there is a smaller range of possible values that the true value is likely to take. There are a number of factors that influence the width of the confidence interval. Larger samples will have narrower confidence intervals and smaller samples are likely to have wider confidence intervals.

Confidence intervals can be calculated for many different types of data, including means, a single proportion, a change in mean or a change in proportion. Examples for different types are data are provided below (adapted from Kirkland, 2012):

Single proportion: Adherence rates for hand hygiene were 83% (95% CI 82 to 84%) in 2007.

A change in proportions: Hand hygiene adherence increased from 83% in 2007 to 94% in 2008, an increase of 11% (95% CI: 10% to 12%).

Options for calculating a 95% CI include use of a statistical package such as Statistical Package for Social Sciences (SPSS), R, Graph Pad, Stata or SAS or online calculators, such as http://vassarstats.net/ (to navigate to CIs, go to proportions or t-tests [for normally distributed continuous data] in left hand column).

5.2 Statistical tests

Statistical tests are used when performing comparisons or testing hypotheses. Statistical tests generate a p value. It is recommended to report p values with the accompanying confidence interval (see Section 2.6 for more information on CI). A p value by itself provides no information about the direction or strength of the relationship and does not allow the reader to judge the clinical relevance of the finding. A p value:

- Explains whether the difference observed in a study is due to chance.
- Measures the strength of evidence against the null hypothesis.
- Is the probability of finding the observed result, or a more extreme one, when the null hypothesis of a study question is true.

Commonly a p value of <0.05 is considered 'statistically significant', meaning the result is unlikely to be due to chance.

Important

It is widely accepted that the p value is overused and misapplied. P values should be reported with a confidence interval.

Although a p value of <0.05 is the standard level of statistical significance this is not an absolute rule. It is an arbitrary value and not intended as a substitute for scientific reasoning.

Having a 'statistically significant' result does not mean that it is 'clinically significant.'

P values are influenced by sample size. A larger sample size will reduce the p value.

Weblinks to source information

https://www.amstat.org/asa/files/pdfs/p-valuestatement.pdf

https://www.nature.com/news/statisticians-issue-warning-over-misuse-of-p-values-1.19503

The statistical test required is dependent on a number of factors, including:

- The hypothesis or question to be addressed by the project.
- The scale of measurement (continuous, ordinal, nominal).
- Data distribution (normal, skewed).
- Paired or unpaired study designs (see Table 4).
- Number of time points.

Each statistical test has assumptions or qualities that must be met in order to use it (see Section 5.3 for more information).

At times simple analysis is appropriate, however there are more complex methods that can adjust for confounders (see Section 4.6 for more information on confounders) or account for clusters/levels within the data (e.g. participants on the same ward). It is advisable to consult a statistician prior to data collection to determine the most suitable approach to analysis.

Analyses for repeated surveys

The best design for measuring the effect of an implemented change is an experimental design with a control group (e.g. randomised controlled trial) because this helps account for known and unknown confounders. This is not always feasible or ethical. Repeated surveys are a form of quasi-experimental design, with inherent limitations for analysis and interpretation (see Section 6 for more information about interpretation). There are two main design approaches for repeated surveys:

Before-after design (audit cycle)

This design compares outcomes before and after a change in practice.

Table 4: Types of statistical tests for within or between-group comparisons where data are collected across two time points only

Data type	**Paired[P] or Independent data[I]**	**Examples of statistical tests**
Continuous data, normally distributed	Independent	Independent t test
Continuous data, normally distributed	Paired	Paired t test
Continuous data, skewed distribution [T]	Independent	Mann Whitney U / Wilcoxon rank sum test
Continuous data, skewed distribution [T]	Paired	Wilcoxon signed rank test
Ordinal	Independent	Mann Whitney U / Wilcoxon rank sum test
Ordinal	Paired	Wilcoxon signed rank test
Nominal	Independent	Chi-square test or Fisher's Exact Test
Nominal	Paired	McNemar's Test

[P] Paired data = data for the same participant before and after implementation of change

[I] Independent = data from different participants before and after implementation of change

[T] Data may be able to be transformed to represent a normal distribution. This requires input from a statistician

It is important to note that comparing data over two time points does not indicate that the change shown is due to the intervention. Change may have occurred with or without the intervention and there may be another explanation(s) for the change. For example, changes during the same time period in staff, exposure to other education, differences between participant groups, seasonality,confounding factors etc. See below for further details 'Interpretation of statistical tests: changes between two time points.'

Repeated measures analysis

An audit cycle may consist of several audits. There are many statistical methods examining three or more time points including repeated measures analysis of variance (ANOVA) or mixed models. The method chosen depends on many factors including data distribution (normal, skewed) and the scale of the measurement (continuous, ordinal, nominal). One of the main considerations is if same participants are followed at each time point or if a new group of participants are measured at each audit. Consulting with a statistician is beneficial when considering repeated measures analysis.

Time series analysis

A method to consider when collecting data over time is time series analysis. A time series design involves collecting data at multiple time points before and after an intervention or policy change. It differs from a simple 'before-after' design (with only two time points) as there are multiple measurements divided over time periods (before and after the implementation of change) It is a design that can measure improvement over time without the need for a control group (Andersson Hagiwara, 2016).

There are several analysis approaches for time series analysis, including statistical process control and interrupted time series with segmented regression. Both methods have the same data requirements, where data are measured over multiple time points. Options vary according to the variable type, sample size etc. An example of graphical display is a run chart or control chart ('u chart') (Figure 6).

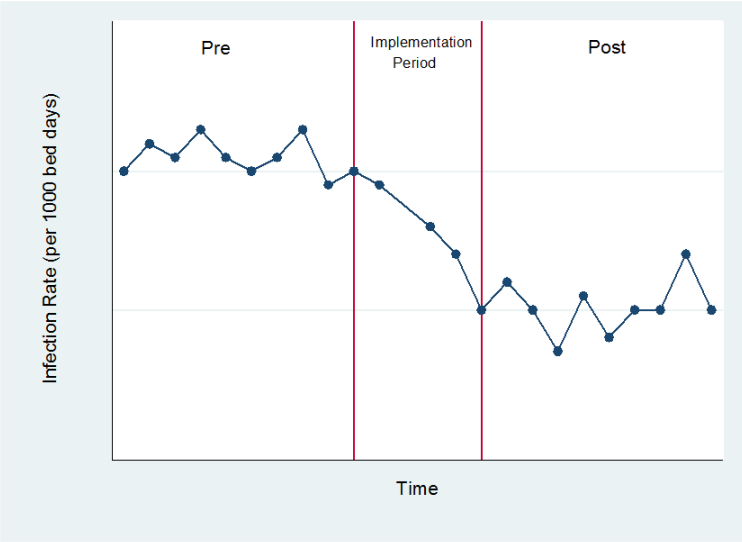


Figure 6: Patient satisfaction using a time series approach before and after implementing a quality improvement change

The statistical process control method is easy to use and requires little statistical knowledge. It is therefore favoured due to ease of interpretation of findings. (Fretheim, 2015; Andersson Hagiwara, 2016). The statistical process control method primarily involves a graphical display of the data. Most simply this involves a run chart, which is a graphical display of the data over time (Figure 6). An extension of the run chart is the Shewhart chart (Andersson Hagiwara, 2016). The Shewhart chart includes additional information, typically a centre line representing the mean or median of the data plus an upper and lower limit, typically 3 standard deviations from the centre line (Andersson Hagiwara, 2016).

The interrupted time series approach using segmented regression has more complex statistical methods than statistical process control but is considered more robust and precise. This analysis examines trends in the baseline period, and then examines shifts in the trend following the implementation of a change. Table 5 compares interrupted time series with segmented regression and statistical process control (Andersson Hagiwara, 2016).

The interrupted time series design is considered one of the strongest non-randomised (quasi-experimental) designs (Harris, 2006; Penfold, 2013) and is commonly used for evaluating the impact of policy change or quality improvement programs (Wagner, 2002). This approach is suitable where a randomised controlled design is not feasible, or where there is limited control over the implementation of a new policy/program. When analysing interrupted time series data using segmented regression it can adjust for expected time trends, however it cannot adjust for the many other factors that can contribute to change (e.g. known and unknown confounders, staff changes, other education exposure, see Section 2.5 for more information on study design).

Features of the time series approach include:

- Frequent use of population level (aggregated) data (e.g. infection rates, number of hospital admissions), however, it can involve individual level data (such as clinical outcomes).
- No need for a control group, with the baseline (before) data acting as the control.
- Ability to test for change in the outcome before and after the implementation of a policy/program, using segmented regression analysis.
- Requirement for a minimum of 8 to 10 time points before and after the intervention (Penfold, 2013).

There are a several statistical packages that can analyse time series data:

- Bernal et al (2017) offers supplementary data and coding files for Stata and R software for segmented regression analysis.
- The itsa command is available through Stata to perform interrupted time-series analysis (Linden, 2015).

Table 5: Strengths and weaknesses of segmental regression and statistical process control (adapted from Andersson Hagiwara, 2016)

Methods	**Use**	**Strengths**	**Weaknesses**
Segmented regression	Assesses strengths of intervention effect over time	Accounts for autocorrelations Can estimate both size of the effect (at different time points) and change in trend over time	Complex analysis Needs a large data set
Statistical process control	Useful for controlling a process and detecting changes in a process	Easy to use without access to advanced software Less need for a large data set	Does not account for autocorrelations Does not measure strength of intervention effect

5.3 Additional resources

Basic data analysis - weblinks

A guide to using data for health care quality improvement (The Victorian Quality Council, 2008)
https://www.aci.health.nsw.gov.au/__data/assets/pdf_file/0006/273336/vqc-guide-to-using-data.pdf

An introduction to analysing quality improvement and assurance data (Healthcare Quality Improvement Partnership, 2018)
https://www.hqip.org.uk/resource/an-introduction-to-statistics-for-local-clinical-audit-and-improvement/#.XA4d2csUlaQ

Choosing the correct statistical test in SAS, Stata, SPSS and R (Institute for Digital Research and Education, University of California, Los Angeles, USA)
https://stats.idre.ucla.edu/other/mult-pkg/whatstat/

Choosing the right elementary statistical test (University of Queensland)
https://my.uq.edu.au/files/10990/statistical-test-table.pdf

Measures of central tendency, testing assumptions and choosing tests (Laerd Statistics)
https://statistics.laerd.com/features-assumptions.php

https://statistics.laerd.com/premium/sts/index.php (fee applies)

https://statistics.laerd.com/statistical-guides/measures-central-tendency-mean-mode-median.php

https://statistics.laerd.com/statistical-guides/measures-central-tendency-mean-mode-median-faqs.php

Quantitative analysis guide: choose statistical test for 1 dependent variable (New York University)
https://guides.nyu.edu/quant/choose_test_1DV

Study design and choosing a statistical test (BMJ)
https://www.bmj.com/about-bmj/resources-readers/publications/statistics-square-one/13-study-design-and-choosing-statisti

What statistical analysis should I use? Statistical analyses using Stata (Institute for Digital Research and Education, University of California, Los Angeles, USA)
https://stats.idre.ucla.edu/stata/whatstat/what-statistical-analysis-should-i-usestatistical-analyses-using-stata/

Basic data analysis

Types of data & measurement scales: nominal, ordinal, interval and ratio (My Market Research Methods, 2019)
https://www.mymarketresearchmethods.com/types-of-data-nominal-ordinal-interval-ratio/

What is the difference between ordinal, interval and ratio variables? Why should I care? (GraphPad, 2009)
https://www.graphpad.com/support/faq/what-is-the-difference-between-ordinal-interval-and-ratio-variables-why-should-i-care/

Basic data analysis – peer-reviews papers

du Prel, J.B., Röhrig, B., Hommel, G., Blettner, M. (2010) Choosing statistical tests: part 12 of a series on evaluation of scientific publications. *Dtsch Arztebl Int.* 107(19):343-8.

Parab, S., Bhalerao, S., (2010) Choosing statistical test. *Int J Ayurveda Res.* 1(3):187-91. doi: 10.4103/0974-7788.72494.

More advanced data analysis - weblinks

ANOVA with Repeated Measures using SPSS Statistics (Laerd Statistics)
https://statistics.laerd.com/spss-tutorials/one-way-anova-repeated-measures-using-spss-statistics.php

Repeated measures analysis with R (Institute for Digital Research and Education, University of California, Los Angeles, USA)
https://stats.idre.ucla.edu/r/seminars/repeated-measures-analysis-with-r/

Repeated measures analysis with SAS (Institute for Digital Research and Education, University of California, Los Angeles, USA)
https://stats.idre.ucla.edu/sas/seminars/sas-repeatedmeasures/

Repeated measures analysis with SPSS (Institute for Digital Research and Education, University of California, Los Angeles, USA)
https://stats.idre.ucla.edu/spss/seminars/repeated-measures/

Repeated measures analysis with Stata (Institute for Digital Research and Education, University of California, Los Angeles, USA)
https://stats.idre.ucla.edu/stata/seminars/repeated-measures-analysis-with-stata/

More advanced data analysis – peer-reviewed papers

Hemming, K., Haines, T.P., Chilton, P.J., Girling, A.J., Lilford, R.J. (2015). The stepped wedge cluster randomised trial: rationale, design, analysis, and reporting. *BMJ. 350*: h391.

Weblinks to The Research Education Program content

Data collection and management (CAHS, Australia, 2019)
https://cahs.health.wa.gov.au/Research/For-researchers/Research-Education-Program/Past-seminars

Section 6: Interpretation of statistical tests and data presentation

This section explains how limitations of cross-sectional study designs and statistical tests affect the reporting of results and their implications. Any audit or baseline evaluation uses observational data. Even after using statistical methods such as adjusting for confounders, sampling bias remains. Making inferences about treatment strategies is discouraged and any findings need to be considered hypothesis generating only (Prasad, 2013). Hypotheses generated from quality improvement projects can provide baseline data and serve as preliminary findings for designing subsequent trials and future grant submissions (Harvin, 2018).

6.1 Statistical significance vs. clinical significance

P values alone cannot determine that a meaningful change has occurred. A statistical difference or association does not infer clinical significance.

What constitutes a clinically significant effect is not always known and there is not always agreement amongst clinicians. Clinical significance relates to how large the effect is, and if this is large enough to consider changes in clinical care.

Refer to CIs when determining clinical significance and consider whether or not the values within the CI are large enough to care about (Skelley, 2011; Sainani 2012).

6.2 Changes between two time points

Comparing data over two time points does not indicate that the change shown is due to the intervention or change in practice for the reasons outlined previously (see Section 1.2 for more information about cause and effect).

Consider the following example. Infection rates were observed over a 12-month period. A policy and training program aiming to improve hand hygiene was implemented and the infection rates observed again during the following 12 months. Figure 7 displays the change in infection rate before and after the training program using a box plot which shows a reduction in infection rates. There is a ‘statistically significant’ difference between the two rates.

However, using these same data, when examining the trend in infection rates on a month to month basis (Figure 8) the rate was reducing over time, regardless of the intervention. After controlling for trend there would not be a significant finding from before and after the intervention.

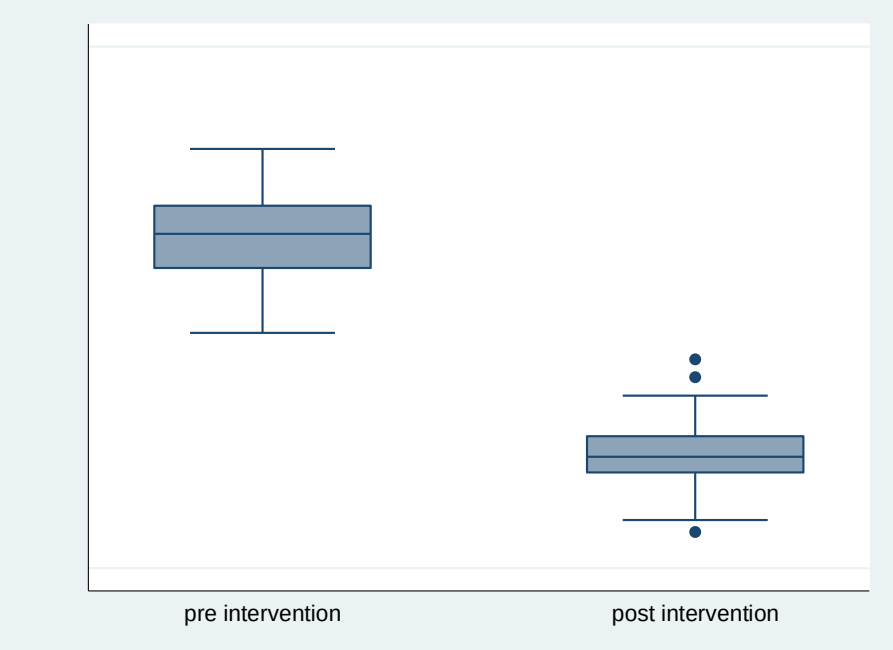


Figure 7: Data presented as box and whisker plots

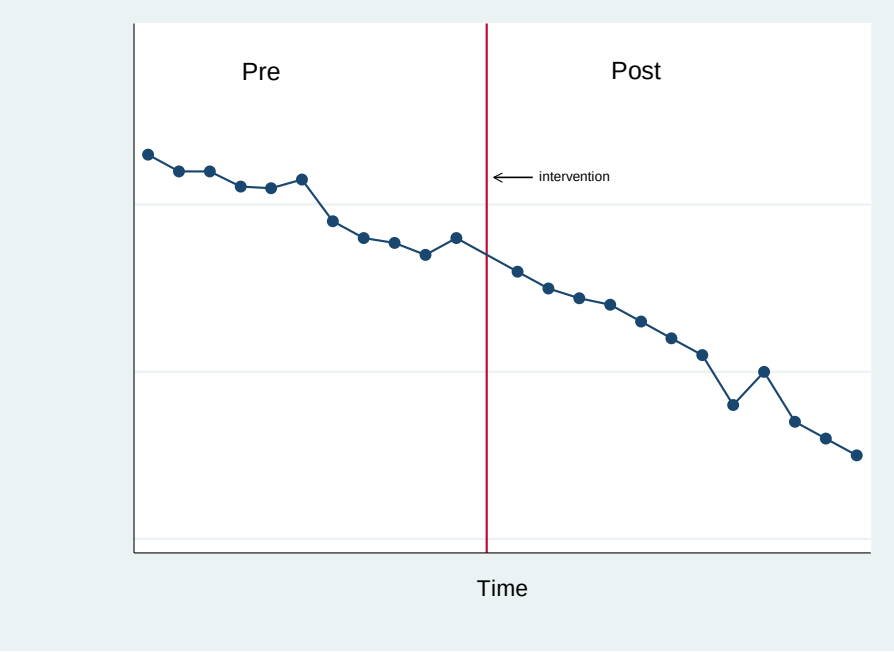


Figure 8: Data presented as a scatter and line plot

When interpreting any changes by comparing data collected across two time points be cautious when there is no control group, or there are not repeated measures. The changes seen may not be attributed the intervention as there may be many other explanations for the change. Time trends are just one. Many unknown confounding factors could have occurred during the same time period that could explain a change, such as staffing changes, other educational influences, changes in funding, equipment, national or state level advertising campaigns etc.

6.3 Causation and association

Interpret data with caution when examining subgroups or associations between variables. Two variables can be associated, but this does not mean that there is a causal relationship. Remember that in a cross-sectional study, all variables are measured at the same time, so it is not possible to infer 'cause and effect' (Prasad, 2013; Bolland 2016).

Consider the following example: 'Suppose we observe that people who drink more than 4 cups of coffee daily have a decreased chance of developing skin cancer. This does not necessarily mean that coffee confers resistance to cancer; one alternative explanation would be that people who drink a lot of coffee work indoors for long hours and thus have little exposure to the sun, a known risk.' It would be necessary to undertake a project with an analytical study design (e.g. case-control, prospective cohort study etc) in order to test this hypothesis.

Be cautious when interpreting the findings for any quality improvement project, and do not overstate your findings. It is enough to describe a change in adherence ***without attributing the change to a particular cause***. The study design cannot answer such a question.

Weblink to source information

https://www.nature.com/articles/nmeth.3587

6.4 Data presentation

This section provides information to assist with presenting the data collected during the audit.

Overview factors to consider when presenting data

Present data objectively and in a logical order. This is usually:

- The study sample, including recruitment and consent rates, completion and drop-outs.
- General characteristics / demographics of participants (e.g. age, gender, location etc).
- Main results related to the primary question, then sensitivity analyses (e.g. key findings stratified by age, gender, postcode etc).
- Secondary findings.

Some basics when describing data:

- Be consistent with presentation, including the number of decimal points for percentages, confidence intervals, as well as tense and use of active voice.
- Use words for numbers when less than 10.
- Use digits for numbers less than one that begin with a zero e.g. 0.3.
- Avoid percentages if sample size <20.
- Do not imply greater precision than your instrument by using too many decimal points.

6.5 Graphs and figures

Graphs and figures are useful for distilling complex information and can reduce the amount of text required to describe results. Do not repeat everything in the text, just the most important points.

- Avoid the use of chart 'gimmicks' such as 3D effects that reduce ease of processing information, as well as difficult colours or fonts.
- Each graph or figure should have a 'stand alone' title describing the study, time period and purpose of the figure or chart.

Examples of reporting results using figures

Bar charts

Bars are divided into discrete categories and are used to compare across groups or categories (Figure 9). These are easy to read and make comparisons between groups accurate and simple. Ideally put bars in rank order of size for improved pre-digestion of information (or alphabetical order, or whatever order makes most sense).

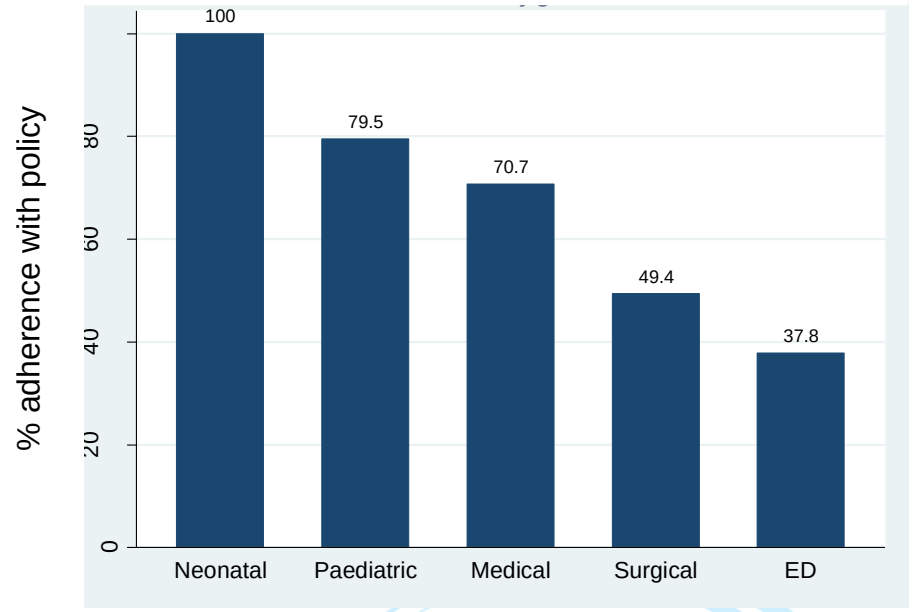


Figure 9: Data presented as a bar chart

Histograms

Histograms are also known as frequency distribution graphs. They are used for continuous data and display the frequency or proportion. It allows you to see the shape or distribution of the data. It can look similar to a bar graph. Bar graphs are used to compare variables, where histograms show the distribution of variables (Figure 10).

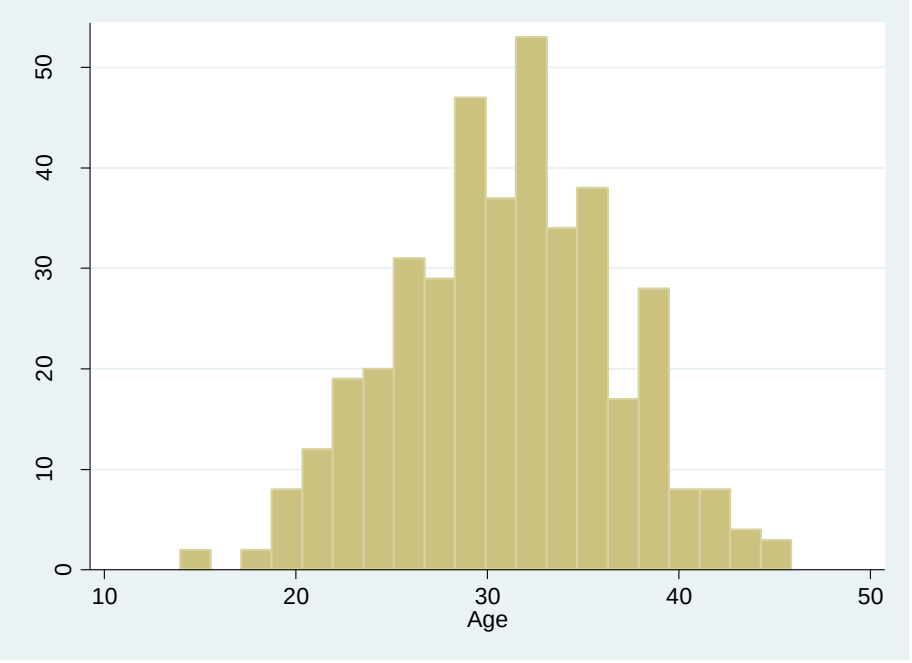


Figure 10: Data presented as a histogram

Line graphs

Line graphs are useful to show a trend over time and display the relationship between two variables (Figure 11).

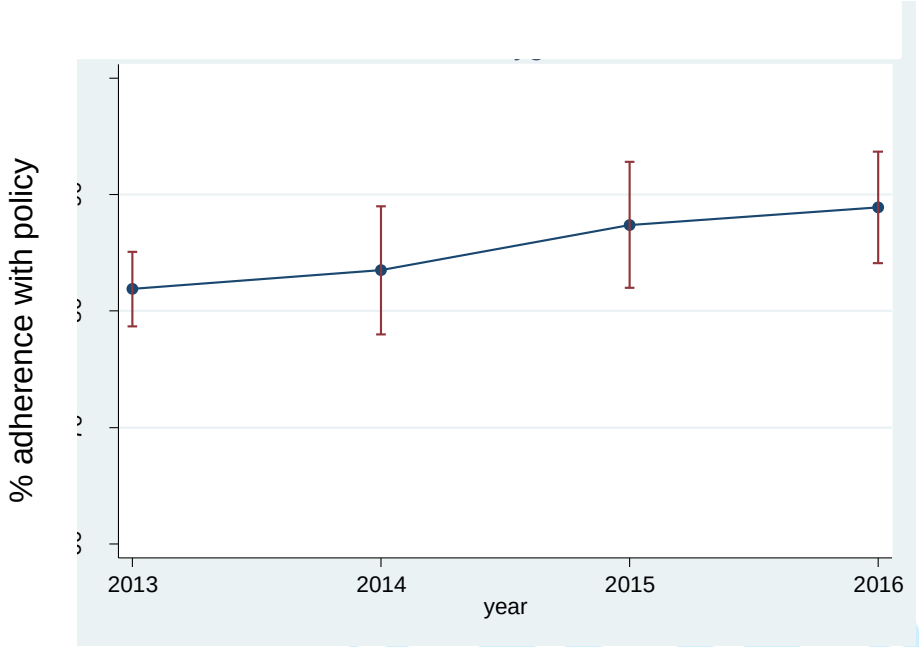


Figure 11: Data presented as a line plot with confidence intervals

Scatter plot

Like a line graph, a scatter plot displays the relationship between two variables. For a scatter plot, dots represent each participant or data point (Figure 12).

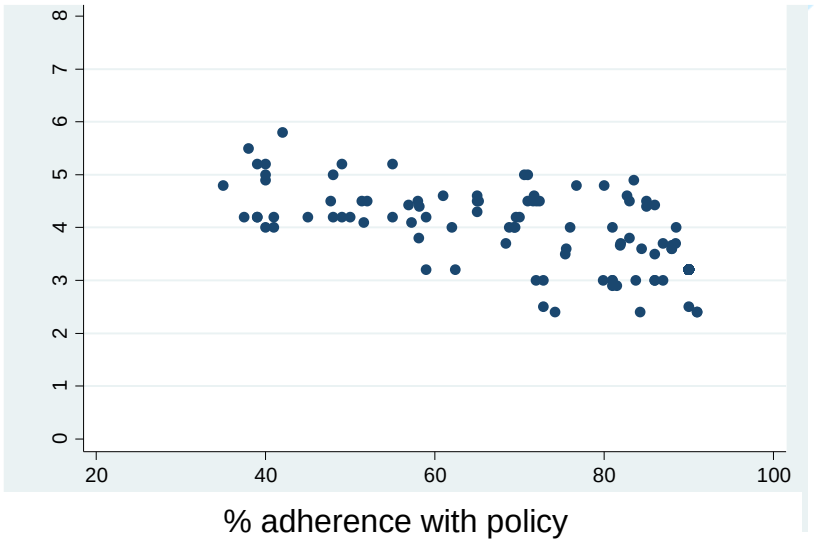


Figure 12: Data presented as a scatter plot

Box and whisker plot

Box and whisker plots (Figure 13) display a statistical summary of a variable, including:

- The median or 50th percentile (line in the centre of the box).
- The 25th percentile (bottom of the box).
- The 75th percentile (top of the box).
- Outliers.
- The distribution of the data (normal or skewed).
- The degree of data dispersion.

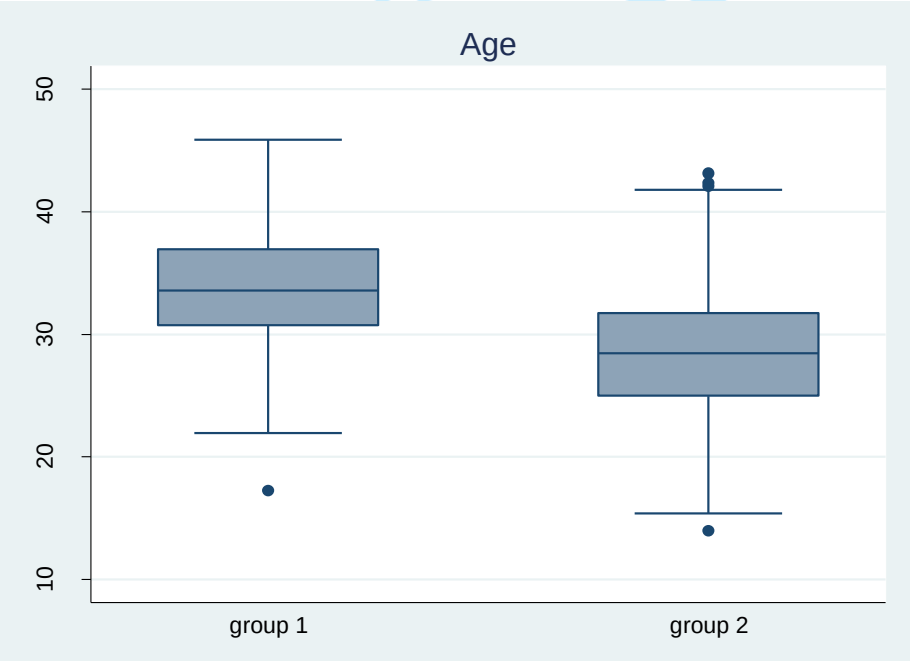


Figure 13: Data presented as a box and whisker plots

6.6 Discussion

The discussion of survey results involves ***critical interpretation***. This should involve other key members of your project team. This is not the place for results to appear for the first time, or where to repeat results. The order of discussion is important. Each key finding can be presented as a paragraph, and is generally structured around:

- What did the study find?
- Were the findings expected or unexpected based on current literature or thinking and why?
- For clinical audits, if benchmarks were not met, what might the reasons be?
- What are strengths and limitations of this particular study/these results?
- What is the impact on current practice/thinking?
- For clinical audits, what changes might be suggested by the results? Are these feasible?
- Recommendations including directions for future study – be specific.

6.7 Additional resources

Presentation of audit as a peer-reviewed publication

https://www.ncbi.nlm.nih.gov/pubmed/31241278

https://www.ncbi.nlm.nih.gov/pubmed/31204842

Presentation of an audit as a poster

http://prehospitalresearch.eu/wp-content/uploads/2015/06/lofaudit-765x1024.jpg

Presentation of audit as a report

https://audit.wa.gov.au/wp-content/uploads/2019/06/Improving-Aborignal-Childrens-Ear-Health.pdf

https://audit.wa.gov.au/wp-content/uploads/2017/08/report2017_14-FSH-w.pdf

Weblinks to The Research Education Program content

Scientific writing (CAHS, Australia, 2019)

Oral presentation (CAHS, Australia, 2019)

https://cahs.health.wa.gov.au/Research/For-researchers/Research-Education-Program/Past-seminars

Section 7: Ensuring knowledge translation

This section emphasises the importance of knowledge translation.

After reporting and interpreting results, it is time to finalise and activate the knowledge translation plan developed at the start of the project (see Section 2.4 for more information on knowledge translation). This should also involve further discussions with key stakeholders to finalise an action plan, and for clinical audits this involves identifying areas for further improvement (Mak, 2019). Potential barriers to change as well as potential enabling factors should be considered. An important first step is to create a table which outlines:

- The required activities.
- Delegation of responsibility for each activity.
- The time frame for completion of each activity.
- Metrics to confirm each activity has been completed.

Remember the project is not complete until the results and recommendations from the project are acted upon. For clinical audits, ideally recommendations are implemented, and the clinical audit cycle continued, by repeating after an interval of time to ensure gains have been achieved.

7.1 Additional resources

Knowledge translation - weblinks

Knowledge Translation Planning Template (The Hospital for Sick Children, Canada, 2013)
http://www.sickkids.ca/pdfs/Learning/79482-KTPlanningTemplate.pdf

Using Change Concepts for Improvement (Institute for Healthcare Improvement, 2019)
http://www.ihi.org/resources/Pages/Changes/UsingChangeConceptsforImprovement.aspx

Knowledge translation- peer-reviewed papers

Dixon-Woods, M., McNicol, S., Martin, G. (2012). Ten challenges in improving quality in healthcare: lessons from the Health Foundation's programme evaluations and relevant literature. *BMJ Quality & Safety.* 21 (10); 876-884.

Levesque, J., Sutherland, K. (2017). What role does performance information play in securing improvement in healthcare? a conceptual framework for levers of change. *BMJ Open*. 7:e014825.

Weblinks to The Research Education Program content

Knowledge Translation (CAHS, Australia, 2019)
https://cahs.health.wa.gov.au/Research/For-researchers/Research-Education-Program/Past-seminars

Section 8: Support and resources

This section provides information on additional support and resources.

Technical support for study design, approval and implementation can be sought from staff working in ethics and governance, safety and quality (in the case of quality assurance projects), biostatistics and research.

8.1 Teaching support

The CAHS Research Education Program

This free, open-access program, for CAHS and non-CAHS staff and students, includes the Research Skills Seminar Series as well as a number of site-specific workshops throughout the year. These aim to facilitate research projects by busy students and staff from any clinical background (e.g. medical, nursing, allied health, pharmacy, laboratory etc). The Research Skills Seminar Series covers key topics across the whole research process. Comprehensive handouts include additional information and resources, along with other known training opportunities. Seminars are available on the day face-to-face or online via hosted video-conference sites or individual devices using the Scopia application and are hosted by the WA Department of Health. Seminars and handouts may be viewed after the day of delivery at https://cahs.health.wa.gov.au/Research/For-researchers/Research-Education-Program/

The Telethon Clinical Research Centre

This centre provides advice to CAHS staff and students on how to access resources related to research (email address: CAHS.TCRC@health.wa.gov.au).

The Western Australian Health Translation Network (WAHTN) Research Education and Training Program:

This program offers on-line training modules on a number of research related topics to staff (free of charge) and students (for a cost) of partner organisations. Certification is provided on completion. Visit: https://retprogram.org/

The Biometrics Team at the Telethon Kids Institute

This team provides fee-for-service consultation services in all aspects of biostatistical and data management support.

Universities

Most universities will offer statistical software and some research training for their students. The most common software packages are described below.

8.2 Database packages

Survey data need to be organised and stored in a system that allows accurate data collection and analysis. Securely storing survey data is essential to protect confidentiality and privacy. The choice of which database package to use is influenced by the style of survey, availability of the database, data storage location (i.e. server or cloud based), data security, data ownership and institute policy. Please refer to individual institution policy on data ownership and storage for further details (see Section 8.4 for information on health policy in WA and further information on data collection and management including naming conventions, meta-data and data management plans).

Excel

Use of Excel is not recommended. Excel is a spreadsheet application and whilst it can be used to store data and is free to use, unlike REDCap and Qualtrics, Excel is not a database management system. It lacks data integrity, accessibility and can be problematic when collecting large amounts of data. Excel does not allow multiple simultaneous users, and version control can therefore be problematic. Columns can be easily re-sorted or erased, leading to major errors. For surveys and data collection a database management system such as REDCap or Qualtrics is preferred to a spreadsheet.

Qualtrics

Qualtrics is an online survey tool. It includes comprehensive survey resources that allow survey design, distribution, analysis and reporting. Qualtrics is available through many universities including Edith Cowan University, Curtin University and the University of Western Australia. Data is stored in the cloud, which can pose policy and ethical issues, especially for surveys where participant consent is not required.

Resources for Qualtrics

https://www.qualtrics.com/au/academia/

https://www.qualtrics.com/resources/

REDCap

REDCap is a secure survey and database web application. It enables building and managing online surveys and databases. It includes comprehensive survey resources that allow survey design, distribution, analysis and reporting. REDCap allows simultaneous data entry and access can be set at various levels for different sites/users. REDCap can be stored on local servers. For example, the Telethon Kids Institute hosts its own instance of REDCap with data captured stored securely on the Institute's onsite servers located within Perth Children's Hospital (PCH) and is protected by a multilayer security protocol.

REDCap is available through many universities and research facilities such as Telethon Kids Institute, University of Western Australia and Curtin University. A number of other sites within WA Health also use REDCap.

Resources for REDCap

The CAHS Research Education Program with Telethon Kids Institute now offers workshops in basic, intermediate and advanced REDCap format throughout the year. Resources related to these can be found at:
https://projectredcap.org/resources/videos/

The WA Health Instance of REDCap is available at: https://cahs.health.wa.gov.au/Research/For-researchers/Research-Education-Program/Additional-resources

Survey Monkey

Survey Monkey is an online survey tool. There are several versions available with various pricing plans, features and security arrangements. Survey Monkey acts as the data custodian for any data entered into their system (https://www.surveymonkey.com/mp/legal/security/). Data are stored in the cloud, which can pose policy and ethical issues. However limited versions are supported by WA Health (e.g. Survey Monkey Enterprise in very specific circumstances).

8.3 Statistical software for data analysis

Graph Pad

This software is easy to learn and use. It is menu driven with drop down boxes, excellent graphics that are easy to edit/modify, and an easy to use help menu. Pricing is reasonable. It has less statistical capabilities compared to other packages, which can be problematic if for example, regression techniques are required.

Resources for Graph Pad

Use of the help menu within the software package is recommended.

SAS

SAS can handle large datasets and is capable of complex analyses. The learning curve and breadth of knowledge is very large. It takes a while for a user to become proficient with SAS.

Resources for SAS

SAS (Institute for Digital Research and Education, University of California, Los Angeles, USA)
https://stats.idre.ucla.edu/sas/

SPSS

This package is commonly taught in universities and is often used by researchers and clinicians. It is easy to use with an intuitive method to accessing menus.

Cost often prohibits its use for those who are unable to access the student licence (annual fee, no perpetual licence arrangement). The basic SPSS package does not perform some commonly used techniques (such as logistic regression).

Resources for SPSS

Laerd Statistics
https://statistics.laerd.com/features-tests.php

SPSS (Institute for Digital Research and Education, University of California, Los Angeles, USA)
https://stats.idre.ucla.edu/spss/

Stata

Stata has user friendly drop-down boxes (menu driven) and syntax. It is relatively inexpensive compared to other packages. Licence durations include 6 months, 12 months or perpetual periods. Stata only holds one dataset at a time and some versions are unable to handle very large datasets. It requires a moderate effort to learn.

Resources for Stata

Dedicated YouTube channel
https://www.youtube.com/channel/UCVk4G4nEtBS4tLOyHqustDA

Online training
https://www.surveydesign.com.au/training.html

Stata (Institute for Digital Research and Education, University of California, Los Angeles, USA)
https://stats.idre.ucla.edu/stata/

Stata Purchasing (Australia)
https://www.surveydesign.com.au/

Resources for learning Stata
https://www.stata.com/links/resources-for-learning-stata/

R package

A free, open source package with a large online community, R is capable of handling large datasets and complex analyses. It has excellent graphical features. It can be difficult to use for those without programming experience.

Resources for R package

R (Institute for Digital Research and Education, University of California, Los Angeles, USA)
https://stats.idre.ucla.edu/r/
There are free on-line classes.

8.4 Additional resources

Comparison of resources

Gathering data using REDCap and Qualtrics (College of Education and Human Services, Utah State University, USA)
https://cehs.usu.edu/research/gathering-data

Quantitative Analysis Guide: Which Statistical Software to Use? (New York University, USA)
https://guides.nyu.edu/quant/statsoft

Health policy documentation for Western Australia

https://ww2.health.wa.gov.au/About-us/Policy-frameworks

https://ww2.health.wa.gov.au/About-us/Policy-frameworks/Information-Management

https://ww2.health.wa.gov.au/About-us/Policy-frameworks/Information-Management/Mandatory-requirements/Collection/Data-Collection-Policy

https://ww2.health.wa.gov.au/en/About-us/Policy-frameworks/Information-Management/Mandatory-requirements/Storage-and-Disposal/Digitisation-and-Disposal-of-Patient-Records-Policy

https://ww2.health.wa.gov.au/About-us/Policy-frameworks/Information-Management/Mandatory-requirements/Governance/Data-Stewardship-and-Custodianship-Policy

https://ww2.health.wa.gov.au/About-us/Policy-frameworks/Information-Management/Mandatory-requirements/Storage-and-Disposal/Patient-Information-Retention-and-Disposal-Schedule-Policy

https://ww2.health.wa.gov.au/About-us/Policy-frameworks/Information-Management/Mandatory-requirements/Access-Use-and-Disclosure/Information-Access-Use-and-Disclosure-Policy

https://ww2.health.wa.gov.au/About-us/Policy-frameworks/Supporting-information/Mandatory-requirements/Information-management/Guidelines-for-the-Release-of-Data

Data management and sharing

Australian National Data Service
https://www.ands.org.au/guides

Weblink to The Research Education Program content

Data collection and management (CAHS, Australia, 2019)
https://cahs.health.wa.gov.au/Research/For-researchers/Research-Education-Program/Past-seminars

Section 9: References

Andersson Hagiwara, M., Andersson Gäre, B., Elg, M. (2016). Interrupted time series versus statistical process control in quality improvement projects. *J Nurs Care Qual*. 31(1): E1-8.

Benjamin, A. (2008). Audit: how to do it in practice. *BMJ. 336*(7655): 1241-1245.

Bernal, J. L., Cummins, S., Gasparrini, A. (2017). Interrupted time series regression for the evaluation of public health interventions: a tutorial. *Int J Epidemiol. 46*(1): 348-355.

Bolland, M.J., Avenell, A., Grey, A. (2016). Qualitative research, observational research, and The BMJ. *BMJ.* 352: i1483.

Boyton, P.M., Greenhalgh, T. (2004). Selecting, designing, and developing your questionnaire. *BMJ*. 328: 1312

Fretheim, A., Tomic, O. (2015). Statistical process control and interrupted time series: a golden opportunity for impact evaluation in quality improvement. *BMJ Qual Saf.* 24(12): 748-52.

Harris, A.D., McGregor, J.C., Perencevich, E.N., Furuno, J.P., Zhu, J., Peterson, D.E., Finkelstein, J. (2006). The use and interpretation of quasi-experimental studies in medical informatics. *J Am Med Inform Assoc.* 13(1): 16-23.

Harvin, J.A., Wootton, S.H., Miller, C.C. (2018). Using quality improvement to promote clinical trials of emergency trauma therapies. *JAMA*. 13;320(18): 1855-1856.

Kirkland, K.B., Homa, K.A., Lasky, R.A., Ptak, J.A., Taylor, E.A., Splaine, M.E. (2012). Impact of a hospital-wide hand hygiene initiative on healthcare-associated infections: results of an interrupted time series. *BMJ Qual Saf.* 21(12): 1019-1026.

Linden, A. (2015). Conducting interrupted time-series analysis for single- and multiple-group comparisons. *Stata J.* 15(2): 480–500.

Leung, W-C. (2001). How to design a questionnaire. *BMJ*; 322:0106187

Mak, D., Maticevic, J. (2019). Preparing for Internship: Clinical Audit Handbook. Fremantle, Australia; Notre Dame University Australia.

Penfold, R. B., Zhang, F. (2013). Use of interrupted time series analysis in evaluating health care quality improvements. *Acad Pediatr.* 13(6 Suppl): S38-44.

Prasad, V., Jorgenson, J., Ioannidis, J. P., Cifu, A. (2013). Observational studies often make clinical practice recommendations: an empirical evaluation of authors' attitudes. *J Clin Epidemiol.* 66(4): 361-366.e4

Sainani, K.L. (2012). Clinical versus statistical significance. *PM&R*. 4(6): 442-445.

Skelly, A.C. (2011). Probability, proof, and clinical significance. *Evid Based Spine Care J.* 2(4): 9-11.

Stone, D.H. (1993). How to do it: Design a questionnaire. *BMJ*. 307: 1264-1266.

Wagner, A. K., Soumerai, S. B., Zhang, F., Ross-Degnan, D. (2002). Segmented regression analysis of interrupted time series studies in medication use research. *J Clin Pharm Ther. 27*(4): 299-309.

Contact

Email: ResearchEducationProgram@health.wa.gov.au

Website: https://cahs.health.wa.gov.au/Research/For-researchers/Research-Education-Program

Department of Child Health Research, Child and Adolescent Health Service

Perth Children's Hospital, 15 Hospital Ave, Nedlands, Western Australia 6009

Disclaimer: The Clinical Audit Handbook has been prepared by the Research Education Program, Department of Child Health Research, Child and Adolescent Health Service, Western Australia as a general guide for conducting clinical audits. It is a generic resource that is not intended to be an exhaustive resource for all requirements for conducting clinical audits. Information contained in the Clinical Audit Handbook is current as of the date published on the front cover, and is subject to change at the discretion of the Child and Adolescent Health Service. It is published online for use only as a general guide for researchers undertaking clinical audits. Individuals are required to consult with their organisation's safety and quality, ethics and governance and/or relevant team and to obtain their own advice to ensure their audit complies with any requirements or processes applicable to their audit.

To the fullest extent permitted by law, the Child and Adolescent Health Service disclaim all liability, direct, indirect, consequential (and whether or not arising out of the negligence, default or lack of care of any person) for any loss or damage suffered by any person out of or in connection with any use of the information contained in this Clinical Audit Handbook.

This page has been intentionally left blank

Cover design and photos: Medical Illustrations, Child and Adolescent Health Service
Ref: M19000299