BMJ Open

Understanding and enhancing responses to distress in the construction industry: protocol for a data linkage study

Lisa Wittenhagen [1,2] Christopher Doran [3,4] John Brady,[5] Edward Heffernan,[1,2,6] Carla Meurk[1,2]

For numbered affiliations see end of article.

Correspondence to
Dr Lisa Wittenhagen;
l.wittenhagen@uq.edu.au

ABSTRACT

Introduction Past research has shown that construction industry workers are at an elevated risk of suicide, however, to date, no study has examined in detail the characteristics of individuals who work in the construction industry and experience distress. This research aims to understand the characteristics, including sociodemographic characteristics, health services utilisation and contacts with non-clinical services, of individuals working within the construction industry and who experience distress, and to quantify the costs and benefits of different help-seeking pathways.

Methods and analysis This study is a data linkage study, based on routinely collected administrative data from construction industry organisations and Queensland Health (QH). Expected outcomes include prevalence of distress, articulated to non-clinical construction industry agencies, descriptive findings on characteristics and help-seeking pathways and health economic analysis. Individuals who experienced distress but who did not contact any of the participating construction industry organisations will not be part of this research.

Ethics and dissemination This data linkage study was approved by Human Research Ethics Committee of The University of Queensland (2021/HE001885). Findings will be presented descriptively to describe the cohort as a whole and stratified by key demographic characteristics, and to provide estimates of prevalence of distress, including timing, frequency and type of contacts. In addition, health economic analysis will be undertaken. Dissemination of findings will be undertaken following consultation of all project investigators, construction industry organisations and peer-led interest groups and lived-experience organisations to ensure translation merit of all findings. Results will be published as peer-reviewed journal articles and publicly available reports.

STRENGTHS AND LIMITATIONS OF THIS STUDY

⇒ This project uses a unique method to inform optimal provision of non-medicalised and peer-led models of care.
⇒ The project examines how distress can be identified within novel datasets, that is, those not traditionally associated with suicide research, but which nevertheless may contain evidence of high levels of unmet need.
⇒ Data linkage studies provide a snapshot of phenomena of interest unencumbered by researcher effects.
⇒ Only individuals who experienced distress and contacted one the participating construction industry organisations will be part of this research, meaning that estimates of prevalence will be conservative.
⇒ While data linkage is undertaken with high rigour, there is a chance for errors within the linked data.

frequency of work opportunities, long working hours and is contingent on complex interactions across different trades, contractors as well as operation requirements.[3 4] These pressures have been reported to be associated with physical, psychological as well as financial distress and suicidality. For instance, past research has shown that construction industry workers are at an elevated risk of suicide,[5 6] and an array of recurrent themes, including the inability to obtain steady employment, injury or major illness and relationship issues, played a role in the lives of individuals who died by suicide and were employed in the construction industry.[7] In addition, while previous research has demonstrated that high levels of distress have a profound negative impact on the economy in general,[8] there is a dearth of information on the economic cost of distress in the construction industry. The construction industry workforce is estimated to account for about 9% of the total Australian gross domestic product.[9] Based on lost productivity and absenteeism, the impact of psychological

INTRODUCTION

About 9% of the Australian workforce is employed within the construction industry, making it the third largest workforce within Australia,[1] and employment within this industry is expected to grow by about 81% by 2025.[2] Work within the construction industry is characterised by changes in the security and

distress among employees of the Australian Coal Mining Industry was estimated at \$AU153.8 million in 2015.[10] A study conducted by Doran *et al* used a costing methodology endorsed by the National Occupational Health and Safety Commission to estimate the economic costs associated with suicide and non-fatal suicidal behaviour to the Australian Construction Industry.[11] Although the authors did not measure the cost of distress per se, they estimated the cost of suicide and non-fatal suicidal behaviour at \$AU1.57 billion. In a parallel study, findings by Doran and colleagues suggested that investing in workplace suicide prevention programmes may have a positive economic return of 4.6:1, suggesting that every \$AU1 invested results in benefits (ie, averted costs) of \$AU4.60.[12]

For the purpose of this study, distress for construction industry workers was defined as 'an emotional state in which individuals feel that they are not in control, overwhelmed or are unable to cope'. This definition was developed following a workshop to examine what distress means to those working in the construction industry. Specifically, 35 workshop attendees (28 males) working across a range of roles within the construction industry discussed in six groups what they associated with distress and how they would define it (2021/HE001047). Based on conversations with people working in the industry and organisations servicing this sector, we have identified that construction industry workers in crisis not necessarily turn to health professionals but rather to peer-led organisations within the industry.

This protocol describes a proposed data linkage study that aims to (1) determine the prevalence of distress, including suicidal crisis, experienced by construction industry workers, (2) detail the health and non-medicalised service utilisation of construction workers in the year prior to and the year following of a distress related contact with a construction industry organisation, (3) examine the demographics, health characteristics and outcomes (including deaths) for individuals in the construction industry who make a distress related call to a construction industry agency and (4) to perform health economic analysis to quantify the costs and benefits associated with medicalised and non-medicalised help-seeking pathways. A detailed understanding of the characteristics, including sociodemographic characteristics, health services utilisation and contacts with non-clinical services, of individuals working within the construction industry and who experience distress will provide valuable information regarding the extent of need for services, facilitate an evaluation of changes in help-seeking behaviours and demand over time and contribute to an assessment of the costs and benefits of different models of service for ameliorating distress. This information is vital to a targeted and systematic approach to provide care to those in need.

METHODS AND ANALYSIS
Study design and setting
This protocol describes a data linkage study. Data linkage is the process of combining data from different data sets into a unified repository for analysis. Its purpose is to harness the power of routinely collected data to inform questions concerning populations across different medical and/or non-medical systems. Data linkage is conducted using mathematical and statistical processes to preserve the privacy and anonymity of individuals whose data are used. This data linkage study is set in the state of Queensland (Australia), which has an estimated resident population of about 5.2 million persons (as of 31 December 2020). Outcomes will relate to individuals who had a distress-related contact with Queensland-based construction industry, or affiliated, organisations. Four construction industry organisations were briefed and identified to participate in this project. An application under section 282 of the Queensland Public Health Act 2005 (the Act) to receive health information for the purpose of research was granted in June 2022 (PHA grant 21.1885). As at September 2022, data are pertaining to the study cohort are being prepared by contributing data custodians for secure transfer to the Queensland Statistical Services Branch (SSB) for data linkage and subsequent supply of deidentified content data to researchers for analysis. It is anticipated that analyses for this study will be finalised mid-2023.

Data sets
A list of data sets included in this data linkage project, including a global description for each dataset, is provided in table 1. Data forming the basis for this research project will be provided by four Queensland-based participating construction industry organisations (to protect the organisations' privacy here referred to as Construction Industry Organisation A, B, C and D). All four participating organisations deal in a direct or indirect way with individuals in distress. Construction Industry Organisation A is a suicide prevention peak body with state-wide coverage. Organisation A provides training to individuals in the construction industry with the aim to reduce suicides among construction industry workers, and direct crisis support via individual case management to construction industry employees in distress. Organisation B is a construction industry redundancy trust which can be contacted by construction industry employees who experience financial hardship due to redundancy or other financial burdens. Organisation C is a construction superannuation fund that can deal with requests to access funds related to financial distress and/or on compassionate grounds. Construction Industry Organisation D is a government-owned, self-funded construction industry insurance provider, which provides compensation in the event of work-related injury or illness.

The Queensland Hospital Admitted Patient Data Collection, the Queensland Hospital Non-Admitted Patient Data Collection, the Registries of Births, Deaths and Marriages Cause of death unit record file, the Emergency Data Collection, Queensland Ambulance Service data and Consumer Integrated Mental Health and Addiction Application data all provide unique information

Table 1 Overview of datasets included in the data linkage study

Dataset	Global description of data that will be requested	Data custodian/agency
Queensland Hospital Admitted Patients Data Collection (QHAPDC)	Data on admitted patients within the state of Queensland, both public and licensed private facilities. Content data requested will include demographic and clinical information, diagnoses (ICD-10), and other items associated with health services contact (eg, length of stay).	(proxy) Queensland Health, Statistical Services Branch (SSB)
Queensland Hospital Non-Admitted Patient Data Collection (QHNAPDC)	Data on non-admitted patients within the state of Queensland. Content data requested will include demographic and clinical information, including chronic diseases, offender health and sexual health.	(proxy) Queensland Health, Statistical Services Branch (SSB)
Registries of Births, Deaths and Marriages (RBDMs) Cause of death unit record file (COD URF)	Data on all deaths within Queensland. Content data requested will include the date of death, and the underlying and contributing cause of death codes according to the International Classification of Diseases (ICD).	(proxy) Queensland Health, Statistical Services Branch (SSB)
Emergency Data Collection (EDC)	Records of all presentations to Queensland hospital emergency departments. Content data requested will include information on individual demographics, triage category, mode of presentation, diagnoses (ICD-10), and the end status of each service episode.	Queensland Health, Hospital Access Analysis Team (HAAT)
Queensland Ambulance Service (QAS) data	Records of all QAS attendances. Content data requested will include the nature and outcomes of emergency and non-emergency patient care provided by QAS.	QAS Commissioner/Director General QH (QAS)
Consumer Integrated Mental Health and Addiction Application (CIMHAA)	Information relating to public mental health services contacts within Queensland. Content data requested for this study will include information on demographics and socioeconomic status, presenting problems and referral source, diagnoses, intervention type, admissions under Queensland's Mental Health Act.	Queensland Health, Mental Health Alcohol and Other Drugs Branch (MHAODB)
Construction Industry Organisation A	Data on individuals in case management. Content data requested for this study will include information around suicidality and service provision.	Construction Industry Organisation A
Construction Industry Organisation B	Construction Industry Organisation B is a Building Employees Redundancy Trust. Content data requested for this study will include claim type and reasons for claims.	Construction Industry Organisation B
Construction Industry Organisation C	Construction Industry Organisation C is a construction industry super fund. Content data requested for this study will include claim type and reasons for claims.	Construction Industry Organisation C
Construction Industry Organisation D	Construction Industry Organisation D is a government owned, self-funded construction industry insurance provider. Content data requested for this study will include the claim type and category, as well as the type of injury.	Construction Industry Organisation D

regarding the quantity and nature of health services utilisation as well as information on a range of demographic and socioeconomic variables, timing and (suicide) deaths.

Cohort selection

The cohort will be identified through keyword and category searches of construction industry datasets Construction Industry Organisations A, B, C and D. Records will be extracted for persons who had a distress-related contact with the participating construction industry organisations over the study period 1 January 2018 to 31 December 2020.

Project investigators LW and CM worked collaboratively with IT specialists and data analysts from each organisation to develop methods for cohort identification and data extraction. Data fields, including free-text boxes, will be searched for a range of distress-related terms using a methodology that has previously been used to identify individuals from free text data.[13] This methodology involves working on iterative data extracts to refine keywords and terms that identify distress, focussing on maximising true positive cases. Terms will be seeded based on the definition of distress. For example, these could include: "out of his/her/their control", "overwhelm*", "distress*", "unable to cope", "not coping" etc. (* indicates allowance for word variations). Additionally, terms and keywords will be chosen to identify specific types of distress, including financial hardship, relationship issues and suicidality. Finally, terms indicating the individual has been referred to specific services that assist individuals in crisis will be identified. Identifying details for individuals who are identified via this method (ie, names, ages or dates of birth, addresses, sex) will be securely transferred by participating industry partner organisations to the SSB, an independent third party, for linkage to Queensland statewide data sets.[14]

It is estimated that the cohort will include about 3000 individuals. It is anticipated that a large proportion of the cohort will be men (in 2019–2020, the industry with the highest proportion of men was the construction industry).[15] The proportion of Aboriginal and/or Torres Strait Islander workforce working within the construction industry in Australia is estimated to lie around 7% (according to the 2016 census, 16 163 individuals of Aboriginal or Torres Strait Islander descent worked within the construction industry and the size of Australia's Aboriginal and Torres Strait Islander labour force was 222 639).Therefore, even though not the focus of this research project, it can be anticipated that the proportion of individuals of Aboriginal and/or Torres Strait Islander descent within the cohort will lie around this estimate. According to data released by the labour market information portal,[2] the median age in 2021 of worker within the construction industry was 38 years.
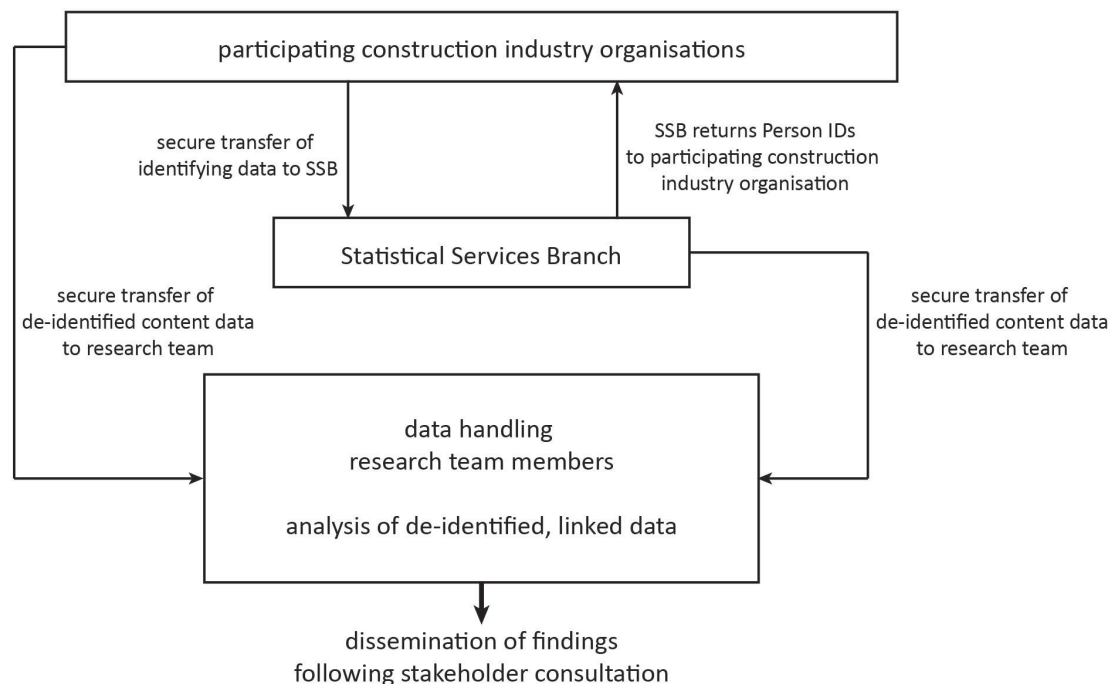
Data linkage

Data linkage will be undertaken in accordance with the Queensland Data Linkage framework,[14] ensuring that the principle of separation is maintained, whereby identifying details are handled separately from content data about an individual. Data will be linked for a period that includes one year either side of the cohort identified (ie, 1 January 2017 to 31 December 2021). Data linkage will involve the transfer of two types of data: (1) identifying data (names, addresses, date of birth, etc) about individuals who have had a past distress-related contact with construction industry organisation A–D, which will be securely transferred to the SSB for the purposes of data linkage and (2) deidentified content data (eg, relevant content data about an individual) which will be securely transferred to researchers at The University of Queensland for the purposes of the study.

One key principle to protect an individual's privacy in data linkage studies is through upholding the separation principle. The separation principle is a mechanism to protect the identities of individuals by separating identifying information from content data of interest. The separation principle ensures that researchers will only have access to deidentified data and at no point in time will they have access to identifying information that could directly identify an individual whose data are being used. This principle works by using an intermediary (ie, the SSB within Queensland Health) to handle identifying information (such as name, address, date of birth) that are required to undertake linkage, and then applying a linkage key which is transferred back to organisation A–D in order for them to supply content data to researchers that are linked at the individual level but fully deidentified (ie, stripped of names and addresses). Linkage of data will be undertaken by the SSB using both probabilistic as well as deterministic linkage techniques as outlined within the Queensland data linkage framework.[14] The SSB will determine whether deterministic or probabilistic linkage techniques will be applied based on the quality of the data received from the participating industry organisations. Based on the established linkage key, content data from participating construction industry organisations will be linked and extracted, including information on marital status, occupation, nature of work and redundancy, financial hardship and total and permanent disability. A schematic of the data transfer and linkage process is provided in figure 1.

Consent

The National Health and Medical Research Council (NHMRC) recognises that there are situations where it is not possible to obtain informed consent from research participants. For this data linkage project, and consistent with the Australian Privacy Principles, a waiver of consent was granted for this project under sections 95 and 95a of the Privacy Act 1988. According to the National Statement on Ethical Conduct in Human Research published by the NHMRC, a waiver of consent can be obtained by researchers for a project 'where it is impracticable to obtain an individual's explicit consent to the use of their information and the purpose of the research cannot be served

Figure 1 Schematic of data transfer and dissemination process. SSB, Statistical Services Branch.

by using non-identifiable information'. The nature of the proposed data linkage study, which will use retrospective data, and the associated impracticalities with obtaining individual informed consent for a large number of individuals, both those whose data are screened for eligibility, and those who are included in the final cohort, means that the researchers believe that a request for a waiver of consent for the release of confidential information for the purposes of research is justified.

Data management

A comprehensive data management plan will reduce risks of disclosure of identity via breaches of data security. Processes of data transfer to SSB of identifying data for the purposes of linkage will be undertaken via secure file transfer protocols (such as Defigo or Sufex). Content data received and generated through the project will be securely transferred to The University of Queensland's Research Data Manager and strictly managed in accordance with The University of Queensland's Research Data Management Policy. This policy ensures that research data are appropriately managed according to recommendations made in The Australian Code for the Responsible Conduct of Research and other relevant Australian legislation. Deidentified data will only be accessible to LW, CD and CM for analysis (who are not members of any of the participating industry organisations and as stipulated on PHA grant 21.1885).

Data will be kept for the prescribed minimum retention period (5 years) and disposed of in accordance with The University of Queensland's Research Data Management Policy. Data will not be made available for other projects unless an amendment for sharing and/or future use of data and/or follow-up research has been approved by

relevant governance bodies. Data are stored on Australian servers and no data are stored overseas.

The project will follow best practice statistical disclosure control techniques, as stipulated by the Australian Bureau of Statistics (Australian Bureau of Statistics, Confidentiality Series, 2017), which include rules for suppressing publication of count or percentage data when numbers are very small. Disclosure control mechanisms protect the identities of individuals from being disclosed, by ensuring that published findings do not inadvertently reveal the identity of any study participants who may possess unusual combinations of characteristics. This can happen in situations where individuals have rare combinations of characteristics (eg, postcode and cultural background) that could make them identifiable.

Data analysis

Descriptive statistics will be used to describe the cohort as a whole and stratified by key demographic characteristics (eg, age, gender, Aboriginal and Torres Strait Islander status, occupational stream matching the Australian and New Zealand Standard Classification of Occupations) as well as health services contacts. In addition, health economic analysis will be undertaken to quantify the costs and benefits of different help-seeking pathways. Data analysis will be conducted in R.[16] Findings will be presented descriptively, to provide estimates of incidents of distress, including timing, frequency and the type of distress-related contacts.

Variables

Variables included (see table 1) in this study will span (but are not limited to) date and time when a distress-related contact was received, demographic information

(including age, gender, ethnicity and geographic locations), information around medical and non-medical services provided (including triage categories, International Classification of Diseases (ICD)-10 codes, type and timing of service provision, service location and facility, referral sources, intervention type, involuntary admissions under Queensland's Mental Health Act, suicide risk, reasons for services provided), timing and nature of deaths (including deaths due to intentional self-harm; ICD-10 codes X60-84)[17]

The main outcome measures will be the frequency and nature of the cohort's (and subgroups within the cohort) engagement with medical and non-medical services before and after the identified index event (ie, distress-related contact with a construction industry organisation) as well as (suicide) deaths. Health services contact will be analysed with a view to examine the frequency of health service use, the type of health service and the service that was provided (including diagnoses). All contacts with a health service will be described in temporal relation to the index event. That is, it will be examined whether a contact preceded and/or followed the index event. Data collected will allow for the calculation of prevalence rate estimates of distress-related contacts and the timing and frequency of contacts with construction industry engagements in times of crises.

Demographic characteristics will be ascribed to an individual probabilistically. Specifically, the age of an individual will be defined as the age of the individual at the first contact with a construction industry organisation (the index event), where age data differ for an individual, the youngest age will be taken. The gender of an individual will be identified as the most common (mode) gender identified for an individual across all data sets. Aboriginal and Torres Strait Islander status is determined by any record showing an individual is of Aboriginal and/or Torres Strait Islander status across all data sets. Plans for analyses may be revised and refined following receipt of the data.

Furthermore, we will quantify the costs and benefits of different help-seeking pathways for individuals who had a distress-related contact with a construction industry agency. As hospitalisation and procedures are associated with a certain cost (and so are non-medicalised or peer-led models of care), health economic analyses can be applied to estimate the costs associated with the different pathways in relation to their benefits and limitations. The health economic analysis will be performed based on information available around timing and health and non-medicalised or peer-led service utilisation available through the linked dataset. All resources associated with healthcare utilisation will be identified, measured and valued to derive an estimate of economic costs. Results will be subjected to sensitivity and uncertainty analyses.

Patient and public involvement

Not applicable

ETHICS AND DISSEMINATION

This data linkage study was approved by Human Research Ethics Committee of The University of Queensland (2021/HE001885). Results will be published as peer-reviewed journal articles and publicly available reports. Dissemination of findings will be undertaken following consultation of all project investigators, construction industry organisations and peer-led interest groups and lived-experience organisations to ensure translation merit of all findings (figure 1).

Author affiliations
[1]School of Public Health, The University of Queensland, Herston, Queensland, Australia
[2]The Park Centre for Mental Health, Queensland Centre for Mental Health Research, Wacol, Queensland, Australia
[3]School of Health, Medical and Applied Sciences, Central Queensland University, Brisbane, Queensland, Australia
[4]Cluster for Resilience and Well-being, Manna Institute, Central Queensland University, Brisbane, Queensland, Australia
[5]Queensland & Northern Territory Branch, MATES in Construction, Brisbane, Queensland, Australia
[6]Forensic Mental Health Service, Metro North Hospital and Health Service, Brisbane, Queensland, Australia

ORCID iDs
Lisa Wittenhagen http://orcid.org/0000-0002-2253-1174
Christopher Doran http://orcid.org/0000-0002-9009-4906

REFERENCES

1 National Skills Commission, Australian Government. Jobs by industry, 2021. Available: https://www.nationalskillscommission.gov.au/jobs-industry-0
2 The Department of Education Skills and Employment, Australian Government. Labour market information portal, 2021.

Available: https://lmip.gov.au/default.aspx?LMIP/GainInsights/IndustryInformation/Construction

3 Ringen K, Seegal J, Englund A. Safety and health in the construction industry. *Annu Rev Public Health* 1995;16:165–88.

4 Gillen M, Baltz D, Gassel M, *et al*. Perceived safety climate, job demands, and coworker support among Union and nonunion injured construction workers. *J Safety Res* 2002;33:33–51.

5 Heller TS, Hawgood JL, Leo DD. Correlates of suicide in building industry workers. *Arch Suicide Res* 2007;11:105–17.

6 Milner A, Niven H, LaMontagne A. Suicide by occupational skill level in the Australian construction industry: data from 2001 to 2010. *Aust N Z J Public Health* 2014;38:281–5.

7 Milner A, Maheen H, Currier D, *et al*. Male suicide among construction workers in Australia: a qualitative analysis of the major stressors precipitating death. *BMC Public Health* 2017;17:584.

8 Hilton MF, Scuffham PA, Vecchio N, *et al*. Using the interaction of mental health symptoms and treatment status to estimate lost employee productivity. *Aust N Z J Psychiatry* 2010;44:151–61.

9 Australian Industry and Skills Committee. Construction, 2021. Available: https://nationalindustryinsights.aisc.net.au/industries/construction

10 Ling R, Kelly B, Considine R, *et al*. The economic impact of psychological distress in the Australian coal mining industry. *J Occup Environ Med* 2016;58:e171–6.

11 M Doran C, Ling R, Milner A. The economic cost of suicide and non-fatal suicidal behaviour in the Australian construction industry. *Int J Ment Health Psychiatry* 2016;02.

12 Doran CM, Ling R, Gullestrup J, *et al*. The impact of a suicide prevention strategy on reducing the economic cost of suicide in the New South Wales construction industry. *Crisis* 2016;37:121–9.

13 Meurk C, Wittenhagen L, Bosley E, *et al*. Suicide crisis calls to emergency services: cohort profile and findings from a data linkage study in Queensland, Australia. *Aust N Z J Psychiatry* 2022;56:144–53.

14 Department of Health. Queensland data linkage framework. Brisbane, QLD, Australia Statistical Analysis and Linkage Unit, Statistical Services Branch, Department of Health; 2021. https://www.health.qld.gov.au/__data/assets/pdf_file/0030/150798/qlddatalinkframework.pdf

15 Australian Bureau of Statistics. Gender indicators, Australia, 2021. Available: https://www.abs.gov.au/statistics/people/people-and-communities/gender-indicators-australia/2020

16 R Core Team. R: a language and environment for statistical computing. Vienna, Austria R Foundation for Statistical Computing; 2022. https://www.R-project.org/

17 World Health Organization. ICD-10 : international statistical classification of diseases and related health problems. Geneva World Health Organization; 2004.