Clinical Epidemiology

Dovepress
open access to scientific and medical research

Open Access Full Text Article

ORIGINAL RESEARCH

The End Rheumatic Heart Disease in Australia Study of Epidemiology (ERASE) Project: data sources, case ascertainment and cohort profile

This article was published in the following Dove Press journal:
Clinical Epidemiology

Clinical Epidemiology downloaded from https://www.dovepress.com/
For personal use only.

Judith M Katzenellenbogen,[1,2,*] Daniela Bond-Smith,[1,*] Rebecca J Seth,[1] Karen Dempsey,[3] Jeffrey Cannon,[2] Lee Nedkoff,[1] Frank M Sanfilippo,[1] Nicholas de Klerk,[1,2] Joe Hung,[1] Elizabeth Geelhoed,[4] Daniel Williamson,[5] Rosemary Wyber,[2,6,7] Anna P Ralph,[3] Dawn Bessarab[1]

On behalf of the ERASE Collaboration Study Group

[1]School of Population and Global Health, The University of Western Australia, Perth, WA, Australia; [2]Group A Streptococcus Research Group, Telethon Kids Institute, Perth, WA, Australia; [3]Global and Tropical Health, Menzies School of Health Research, Charles Darwin University, Darwin, NT, Australia; [4]School of Allied Health, The University of Western Australia, Perth, WA, Australia; [5]Aboriginal and Torres Strait Islander Health Branch, Queensland Health, Brisbane, QLD, Australia; [6]Office of the Chief Scientist, The George Institute for Global Health, Sydney, NSW, Australia; [7]The University of Sydney, Sydney, NSW, Australia

*These authors contributed equally to this work

Purpose: Acute rheumatic fever (ARF) and rheumatic heart disease (RHD) persist as public health issues in developing countries and among disadvantaged communities in high-income countries, with rates in Aboriginal and Torres Strait Islander peoples in Australia among the highest recorded globally. A robust evidence base is critical to support policy recommendations for eliminating RHD, but available data are fragmented and incomplete. The End RHD in Australia: Study of Epidemiology (ERASE) Project aims to provide a comprehensive database of ARF and RHD cases in Australia as a basis for improved monitoring and to assess prevention and treatment strategies. The objective of this paper is to describe the process for case ascertainment and profile of the study cohort.

Patients and methods: The ERASE database has been built using linked administrative data from RHD registers, inpatient hospitalizations, and death registry data from 2001 to 2017 (mid-year). Additional linked datasets are available. The longitudinal nature of the data is harnessed to estimate onset and assess the progression of the disease. To accommodate systematic limitations in diagnostic coding for RHD, hospital-only identified RHD has been determined using a purposefully developed prediction model.

Results: Of 132,053 patients for whom data were received, 42,064 are considered true cases of ARF or RHD in the study period. The patient population under 60 years in the compiled dataset is more than double the number of patients identified in ARF/RHD registers (12,907 versus 5049). Non-registered patients were more likely to be older, non-Indigenous, and at a later disease stage.

Conclusion: The ERASE Project has created an unprecedented linked administrative database on ARF and RHD in Australia. These data provide a critical baseline for efforts to end ARF/RHD in Australia. The methodological work conducted to compile this database resulted in significant improvements in the robustness of epidemiological estimates and entails valuable lessons for ARF/RHD research globally.

Keywords: rheumatic fever, rheumatic heart disease, indigenous, epidemiology, linked data

Introduction

Global Burden And Pathophysiology Of ARF And RHD

In 2018, the 71st World Health Assembly adopted a resolution on rheumatic fever and rheumatic heart disease (RHD), acknowledging them as preventable yet greatly burdensome diseases.[1] Acute rheumatic fever (ARF) is an autoimmune reaction to a Group A Streptococcus (Strep A) infection in the throat or skin.[2,3] ARF, especially when recurring, can cause chronic RHD, characterized by permanent cardiac valve damage resulting in premature morbidity and mortality. The disease trajectory

Correspondence: Judith M Katzenellenbogen
School of Population and Global Health, The University of Western Australia, M431 35 Stirling Highway, Crawley, Perth, WA 6009, Australia
Tel +61 8 6488 1001
Fax +61 8 6488 1188
Email Judith.katzenellenbogen@uwa.edu.au

submit your manuscript | www.dovepress.com
DovePress
http://doi.org/10.2147/CLEP.S224621

© 2019 Katzenellenbogen et al. This work is published and licensed by Dove Medical Press Limited. The full terms of this license are available at https://www.dovepress.com/terms.php and incorporate the Creative Commons Attribution – Non Commercial (unported, v3.0) License (http://creativecommons.org/licenses/by-nc/3.0/). By accessing the work you hereby accept the Terms. Non-commercial uses of the work are permitted without any further permission from Dove Medical Press Limited, provided the work is properly attributed. For permission for commercial use of this work, please see paragraphs 4.2 and 5 of our Terms (https://www.dovepress.com/terms.php).

typically starts in childhood with a median age of 12 years at first ARF diagnosis and RHD commonly diagnosed later in adolescence and early adulthood.[2,4] There is currently no vaccine against Strep A infection and no cure for RHD. Secondary prevention protocols after ARF involve regular, long-acting penicillin injections every 3 to 4 weeks for several years to prevent further Strep A infections and mitigate disease progression. Surgical intervention is often required in the management of severe RHD.[5–7]

It is estimated that RHD affects around 30 million people globally, often leading to permanent disability and 305,000 premature deaths annually.[8] Strep A infections are associated with inadequate hygiene and overcrowding, which are driven by insufficient access to good sanitation and adequate quality housing.[9] Consequently, RHD is considered to be a disease of disadvantage. Because of its significant environmental and socioeconomic etiology, ARF/RHD is an endemic public health problem in many low- and middle-income countries,[5,6,8,10] while the disease has been virtually eradicated from the general populations of high-income countries. However, among disadvantaged communities, the disease remains hyper-endemic. In Australia, RHD predominantly impacts Aboriginal and Torres Strait Islander people (hereafter respectfully referred to as Indigenous).[11–13]

ARF And RHD In Australia: The Need For Baseline Data

Indigenous people in Australia are the nation's first peoples, comprising 3.3% of the population. Indigenous Australians have a younger age structure than the rest of the population (median age 23 vs 38 years[14]) and experience an eight-year shorter life expectancy.[15] Indigenous people in Australia are reported to have some of the highest RHD prevalence rates in the world,[5] with the differential in RHD burden between Indigenous and non-Indigenous Australians being one of the highest of all disease groups reported.[16] These inequities are complex and intersecting, with the underlying cause relating to historical and ongoing colonization that has led to the disruption of culture, dispossession of ancestral lands, forced removal of children from families, and systemic social and economic disadvantage.[17–19]

In 2009, the Australian government introduced the Rheumatic Fever Strategy (RFS), which focussed on building the infrastructure (including RHD registers) and disease control capacity in the four jurisdictions where the disease burden is the highest. These are Northern Territory (NT), Queensland (QLD), South Australia (SA) and Western Australia (WA). New South Wales (NSW) established its own register in 2015. These registers include patient recall systems to support secondary prophylaxis of ARF, care coordination, and support for the health workforce amongst a range of other strategies.[20] An outline of the components of the RFS, contemporary guidelines for clinical management, epidemiological data sources, and ARF and RHD estimates can be found at https://www.rhdaustralia.org.au/arf-rhd-guideline. In 2016, the END RHD Coalition (https://endrhd.org.au/) was formed by research, professional, non-governmental, and Aboriginal and Torres Strait Islander organizations to advocate for a more comprehensive government strategy to eliminate ARF/RHD as a public health concern in Australia by 2031. This movement is gaining momentum.[21,22]

To develop realistic targets for disease control, comprehensive national baseline data to calculate the incidence and prevalence of ARF/RHD across Australia are required, including important stratifications by age, sex, Indigenous status, geography, and other indicators. However, data on ARF/RHD burden in Australia are fragmented and sources operate independently so that comprehensive and accurate data cannot be obtained from a single source. The five jurisdictional RHD registers provide the most detailed data available, but mainly capture Indigenous and children and young adults and generally have incomplete coverage of the ARF/RHD patient population, due to resource and administrative constraints. Primary health care data, which would provide useful information regarding ARF/RHD, are not easily accessible to researchers, and International Classification of Disease (ICD) codes in hospital data have systematic limitations, usually overestimating hospitalizations with RHD.[23]

Filling this knowledge gap is the primary purpose of the End RHD in Australia: Study of Epidemiology (ERASE) Project. The World Health Organization (WHO) RHD resolution explicitly notes the paucity of data as a barrier to progress in addressing ARF/RHD and highlights the need for reliable estimates of the national burden of RHD as a priority activity.[24]

Linked Administrative Data As A Critical Tool For Burden Estimation

The ERASE Project uses linked administrative data, including information from ARF/RHD registers, hospital data, death records, and various other sources (see

Supplementary Table 1), to create a comprehensive database for characterizing the ARF/RHD patient population and estimating the burden of ARF and RHD. Using linked data provides more reliable estimates of the ARF and RHD burden as the linked dataset allows for a person's records to be followed across different data collections, compensating for the incompleteness of data from a single source. The longitudinal nature of the data then allows an accurate estimation of disease onset and progression.

The aims of the ERASE Project are to determine the baseline burden of ARF/RHD in Australia and to develop further insights into the progression of the disease as a basis for improved monitoring and to assess secondary prevention and treatment outcomes. The purpose of this paper is to (1) provide a detailed account of the compilation and preparation of the ERASE database; (2) compare the information available from the different data sources and profile the ARF/RHD patient population; (3) disseminate our data collection and methodological work on ARF/RHD to facilitate collaboration and access to our methods for other researchers.

Materials And Methods

Data were available for the five Australian jurisdictions where the disease burden is the highest and where ARF/RHD registers have been established: NSW, NT, QLD, SA and WA. Together these five jurisdictions are home to 86% of Indigenous Australians (at 30 June 2016).[14]

Description Of Data Sources And Data Linkage

Probabilistic data linkage was undertaken separately by the relevant linkage units in each jurisdiction. De-identified datasets including unique person identifiers (and for hospital records, admission identifiers) were provided to the authors. Within-jurisdiction data linkage was undertaken for WA, QLD, and NSW. SA and NT were linked cross-jurisdictionally, such that an individual could be followed across state boundaries, due to the substantial cross-regional flows of patients between regions. The well-documented geographical high mobility of many Indigenous Australians – particularly those living in remote regions – reflects a range of (often predictable) familial, cultural, logistic and service-seeking imperatives.[25,26] While some movement across borders also applies to the other jurisdictions in Australia, this happens on a smaller scale[27] and the benefit of cross-jurisdictional data was outweighed by the substantially shorter delivery times of jurisdiction-specific data. As with many data linkage projects in Australia,[28,29] receipt of the data occurred more than three years after first initiating the multiple applications for data. This delay reflects one of the major limitations of obtaining health data for surveillance in Australia, exacerbated by the federal structure, where legal frameworks and bureaucratic processes/requirements differ across jurisdictions.[28] Supplementary Table 1 details which content data collections are available for different jurisdictions and provides a brief description of the type of information included. Five data collections are available for all jurisdictions: ARF/RHD registers, inpatient admissions, emergency department presentations, the National Cardiac Surgery Database of the Australian & New Zealand Society of Cardiac & Thoracic Surgeons (ANZSCTS),[30] and death records. However, the ARF/RHD registers have varying establishment dates across jurisdictions and, similarly, cardiac units started contributing data to the ANZSCTS database at different points in time.

Definition Of The Data-Generating Cohort

The data-generating cohort includes all people in NSW, NT, QLD, SA, or WA for whom data were linked by the data linkage units; specifically, any person who has/had:

1. a record on an Australian ARF/RHD register (has been "registered") or
2. at least one hospital admission (including all public and most private hospitals) with an ICD-10-AM (Australian Modification) code of I00 to I02 (ARF) or I05 to I09 (RHD) in any diagnosis field or
3. a death registry record coded as I00 to I02 (ARF) or I05 to I09 (RHD) or an equivalent free text field (for later periods where the cause of death coding was not available) in any cause of death field.

Data time frames vary between jurisdictions (see Supplementary Table 1), but all cover the period between 2001 and 2017 (mid-year), with more recent data available for most jurisdictions. WA hospital and death data are available from 1980.

The data-generating cohort itself or any individual data source alone is not a suitable basis for epidemiological analyses of ARF/RHD. In particular, systematic biases in the predictive accuracy of the ICD codes for RHD (I05-I09) in the hospitalization data have been identified,

limiting their usefulness for case identification and analysis.[23] Methodological work has been conducted by the ERASE Project to strengthen the robustness of ARF and RHD case identification in administrative data.

Identification Of ARF And RHD Cases

Since ARF is a potentially recurrent acute condition, it is important to identify every distinct episode of ARF as accurately as possible. An ARF episode was defined as an ARF record with >90 days free of an ARF diagnosis from any data source.[31] In contrast, RHD is a chronic condition. Hence, the objective was to correctly establish the onset of the disease by identifying the earliest reliable record of RHD for each person. Further, an effort was made to distinguish severe from mild or moderate RHD cases, in accordance with the priority classification operationalized in the Australian RHD guidelines.[32]

A first-ever record of ARF or RHD was ascertained by employing a lookback period where a person's records were searched for any previous record of ARF/RHD across all available data sources. The ARF/RHD status information presented in this paper uses all data from the earliest available. Forthcoming publications as part of the ERASE Project will conduct sensitivity analyses to justify the appropriate choice of lookback periods and/or data source(s) for specific analyses. Methods for identifying "first-ever" status for ARF and RHD will be described in detail in each relevant publication, as these may differ, depending on the research question, study period, lookback period and data sources used. Various sources of validation of RHD status (including ARF/RHD registers, ANZSCTS database, pediatric surgical data, QLD and WA file audits, see Supplementary Table 1), are available where ARF/RHD status has been verified through case follow-up, file audits or availability of detailed clinical variables.

Identification Of ARF And RHD Cases Using ARF/RHD Register Data

According to the Australian RHD guidelines, every case of ARF and RHD should be clinician-notified for inclusion in a jurisdictional ARF/RHD register, if a register is existent in the jurisdiction where the diagnosis occurred.[32] Hence, register data have been used in much of the previous research on ARF and RHD in Australia.[4,33–38] However, there are substantial issues with regard to case capture and data accuracy on these registers, particularly in regions other than the Northern Territory (see Results for details). In addition, echocardiography screening studies in school children in high-risk Northern Australian communities found that 53% of children with definite RHD by World Heart Federation criteria had no previous RHD diagnosis recorded.[39] This suggests substantial under-diagnosis and consequently gaps in register completeness.

An ARF case was defined as any person with at least one episode of ARF recorded on a jurisdictional ARF/RHD register. A person was defined as having RHD from the earliest date of RHD assessment that was evaluated as "mild", "moderate", or "severe" RHD, or surgery for RHD was recorded on a jurisdictional ARF/RHD register. A person's RHD status was determined to be "severe" from the first date that an RHD assessment status was recorded as "severe" or the register had a record of an RHD-related surgery or procedure (Australian Classification of Health Interventions blocks 621–638).[40]

Identification Of ARF And RHD Cases Using Hospital Data

In addition to register records, hospitalization data can be used to identify cases of ARF/RHD, especially since data are collected in a relatively standard manner over time and across jurisdictions. A large proportion of cases are likely to be hospitalized, particularly since the release of the 2012 Australian RHD guidelines, which recommended that every person diagnosed with suspected ARF should be hospitalized upon onset of symptoms as soon as possible for specialist review and confirmation of the diagnosis. People living with RHD usually require hospital care intermittently to treat their symptoms and receive specialist routine-recommended care.

Substantial methodological work was conducted as part of the ERASE Project to develop robust methods for identifying ARF and RHD cases in hospital admission data. Principal diagnosis was used as the primary identifier of acute episodes of ARF. In addition, an extended definition of ARF was developed that implements the Australian diagnostic criteria for ARF in detail. The extended definition considers the secondary diagnosis of ARF if coded alongside a principal diagnosis of a key symptom of ARF (fever, polyarthritis, subcutaneous nodules, heart block/electrical conduction abnormalities, or other heart-related symptoms) and excludes cases for which alternative diagnoses for these symptoms were provided.[32] Details regarding the inclusions and exclusions for this definition can be requested from the authors. This definition provides an upper bound estimate of ARF incidence. Upon implementation of this extended definition, only a small number of

additional episodes were identified (n=35, 0.5% for the analysis cohort). This can be interpreted as providing additional confidence in the reliability of considering only principal diagnoses.

Previously, we identified systematic issues with the ICD codes for RHD.[23] For example, false positives may occur because of the ICD codes for nonspecific valvular heart disease default to RHD (for example, I05.9, I06.9, I07.9, I08.9). In 2016, the ERASE Project consulted clinical researchers, epidemiologists, RHD control staff, and government health analysts to propose a qualitative algorithm that evaluates the reliability of ICD codes for RHD.[23] It was validated on a sample of RHD cases (n=368) from selected tertiary hospitals in WA[41] resulting in a substantially improved positive predictive value for Indigenous patients ≤35 years but less improved for other subgroups.[23] Consequently, a more quantitative approach was undertaken to develop a prediction model for RHD ICD codes based on a large dataset containing validated cases and non-cases from QLD and the NT (n=7555).[42] With this prediction equation, an area under the receiver operating characteristic curve (AUC) of 0.93 could be achieved.[42] This prediction model also alleviates misclassification of ARF as RHD which had not been addressed in the previous literature. Because of the improvement in case ascertainment, the ERASE Project defines a person as an RHD case from the date of the earliest hospital admission date identified to be RHD by this prediction model.

For RHD cases, as defined above, "severe" RHD was determined to commence with the first hospital admission with heart failure (as identified by ICD-10 code I50) or an RHD-related surgery or procedure (Australian Classification of Health Interventions blocks 621–638).[40]

Hospital admission and separation dates were adjusted for intra- or inter-hospital transfers by considering any admission occurring on the same or the next day as their previous separation date or any nested admissions as part of the same hospital episode.

Identification Of RHD Cases Using Surgical Data

Adult Cardiac Surgery Data

The ANZSCTS database provides another opportunity to identify people with RHD, albeit for late-stage cases. A person was defined to be an RHD case if a surgery with a "rheumatic" valve pathology was recorded in the ANZSCTS database. A valve-specific variable available in the database identifies surgeries with a rheumatic valve, but only RHD cases can be identified (RHD non-cases not identifiable), because other possible categories (for example, valve repair) do not exclude a history of RHD.

Paediatric Cardiac Surgery Data

Paediatric patients from SA and NT are transferred interstate to the Melbourne Royal Children's Hospital (Victoria) for surgical procedures for RHD. Data were extracted from the Royal Children's Hospital for paediatric RHD patients from SA and the NT in our cohort.

Identification Of ARF And RHD Cases Using Other Data

Emergency department presentations data were used to identify ARF cases (persons) and episodes (per person). Because of the concerns with RHD diagnosis codes for the inpatient data, RHD diagnoses recorded in the emergency department data were only considered for cases validated through other sources.

Primary care data on ARF/RHD were only available for ~60 government clinics in the NT and were considered for those cases.

Finally, the ERASE Project has access to an additional 885 cases validated through file audit from QLD obtained from file audits as part of a case-finding program run by the QLD RHD control program that was linked to the cohort. In addition, 368 separately validated cases from tertiary hospitals in WA were also linked.[41]

Definition And Characteristics Of The Study Cohorts

Given the plethora of data sources for case identification and their individual strengths and weaknesses, defining study cohorts required careful consideration in order to balance several objectives. These include the following:

- Reliable identification of true positive cases
- Maximizing accurate case capture
- Comparability of case counts over time and across jurisdictions (especially to allow for reliable projections for the high-burden jurisdictions at the quasi-national level)
- Comparability to data used in previous research and data monitored by the ARF/RHD control programs and peak bodies.

Based on these considerations, three study cohorts were defined, henceforth referred to as register, analysis, and expanded cohorts (Table 1). Figure 1 visualizes the cohorts' relationships to each other and to the original

data-generating cohort while Table1 describes the cohorts and their rationale. All cohorts were based on data between 1 July 2001 and 31 December 2017 guided by the availability of linked data across jurisdictions.

Derivation Of Key Demographics

Key demographics, in particular month/year of birth, sex, country of birth, Indigenous status, population category, and area of residence (by region and jurisdiction), were derived and harmonized using information from the ARF/RHD registers and hospital and death records.

Month Of Birth, Sex, And Country Of Birth

For month of birth, sex, and country of birth, the mode of all observed values for a person was chosen within and across the available data collections. This was implemented using the following steps:

1. For hospital data, if multiple records for a person were available, the mode of all available records for a person was determined. If there was an equal number of records of different values, the mode was set to missing.
2. The mode of each was then calculated across the ARF/RHD register, death data, and the mode of the hospital data from step 1. If no mode could be calculated, then preference was given to the ARF/RHD register, followed by the death data, followed by hospital data.

Indigenous Status

In Australia, Indigenous people are under-represented in administrative data, resulting in limitations for national estimates. However, Indigenous status recording has seen marked improvements in reporting in the past 10 years.[43] The under-reporting may be due to non-recording or misclassification and arises for a number of reasons, including the propensity of the person to identify publicly as Indigenous, health service staff not enquiring about Indigenous status, or administrative errors.[43] In previous research, "ever" Indigenous was found to over-count while "all" Indigenous is found to substantially undercount the number of Indigenous people in administrative records.[44,45] For the ERASE Project, exploratory work was undertaken to compare various assignment options with the "Getting our Story Right" (GOSR) indicator that has been developed for WA to maximize the predictive power of Indigenous status assignment using multiple administrative data sources.[46] Detailed information regarding the concordance of our chosen algorithm – conceptually similar to the GOSR model – with the GOSR indicator for WA can be obtained from the authors.

Indigenous status was initially assigned at the data source level. For hospital data, where persons had more

Table 1 Overview Of Definitions And Rationale For The Three Study Cohorts Of The ERASE Project

	Register Cohort	**Analysis Cohort**	**Expanded Cohort**
Definition	Any person with: • an ARF/RHD record on an ARF/RHD register	Any person with: • an ARF/RHD record on an ARF/RHD register or • a hospital admission determined as a ARF/RHD diagnosis or • a surgery with a rheumatic valve recorded in the ANZSCTS database or a record in the paediatric RHD surgery data set for SA/NT cases	Any person with: • an ARF/RHD record on an ARF/RHD register or • a hospital admission determined as a ARF/RHD diagnosis or • a surgery with a rheumatic valve recorded in the ANZSCTS database or a record in the paediatric RHD surgery data set for SA/NT cases • an ARF or RHD (for validated cases only) diagnosis in the emergency department data • an ARF/RHD diagnosis in the available primary health care data or • whose ARF/RHD status has been validated through other sources
Rationale	Comparable to control program and (some) previous research data and reliable validation at the expense of case capture	Balancing reliable case identification with maximising case capture and comparability across jurisdictions and time	Maximising case capture for the available data at the expense of comparability

Abbreviation: ERASE, End RHD in Australia: Study of Epidemiology.

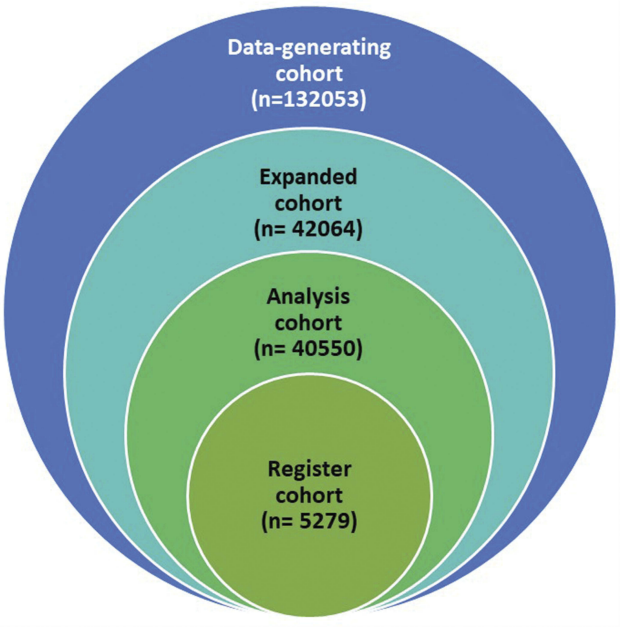


Figure 1 Venn diagram of data-generating and study cohorts.

than two records, people in the study were coded as Indigenous if they were recorded as Indigenous in at least two records. If a person had 2 or fewer records, they were recorded as Indigenous if at least one record was coded as Indigenous. For register, ANZSCTS, death and other data sources, the recorded identifier was assumed correct. Overall, a person was deemed Indigenous for our study if they were flagged to be Indigenous in at least one of the above available data sources.

Population Category

We grouped our cohort into three population categories: Indigenous, immigrant from low-income or lower-middle-income country (ILIC), or other Australian. ARF/RHD is known to be substantially more prevalent in low- and lower-middle-income countries and for some population groups in high-income countries (for example, Maori and Pacific Islanders living in New Zealand/Australia are known to also have an inequitable burden of ARF and RHD).[2,6] Data directly recording population category (for example, Māori/Pacific Islander, available from the NT and SA ARF/RHD registers and the QLD surveillance data) were used, where available. If unavailable, the population category was assigned based on the World Bank Country Income classification status for the financial year 1996 of the person's recorded country of birth.[47] A person was recorded as an "immigrant from a low-income or lower-middle-income country" if the recorded country of birth was a low-income or lower-middle-income country or New Zealand (given likely capture of Māori/Pacific Islander persons in the ARF/RHD-coded data-generating cohort). The remaining patients were identified as "Other Australian".

Region And Jurisdiction Of Residence

Region of residence was represented by Indigenous Regions (IREG), the highest level of aggregation of the Indigenous Structure designated by the 2011 Australian Statistical Geographical Standard (ASGS).[48] These regions have been designed to better reflect the spatial distribution of Indigenous communities than the standard ASGS area definitions. IREG is also the level for which the Australian Bureau of Statistics provides the most reliable sub-jurisdictional population estimates for Indigenous populations by 5-year age groups and sex.

Due to limitations in the availability of the underlying geographical data in the study's data collections and official geographical concordance files, IREGs were assigned based on 2012 postcodes for NT, QLD and WA, on 2011 Statistical Areas Level 2 for NSW and 2014 Localities for SA. If a record could not be matched, matching to previous versions of the respective statistical aggregate was attempted. Since IREGs do not cross state boundaries, jurisdictions could be uniquely assigned based on IREG information.

Results

Data on 132,053 individuals were received from RHD registers, inpatient hospitalizations, and death registry data from five Australian jurisdictions (2001 to 2017, mid-year). Of these, 42,064 are considered to represent true cases of ARF or RHD (Figure 1).

Case Capture In The Different Data Sources

Gaps were found in completeness of the ARF/RHD register data. Table 2 shows the cumulative frequencies of persons by the data source for the expanded cohort (where each person was uniquely counted towards a data source starting with ARF/RHD registers, hospital, etc.). Between 2001 and 2017 for individuals under 60 years of age at the time of their first diagnosis of either ARF or RHD, almost two-thirds of persons with an ARF or RHD diagnosis identified in the expanded cohort (n= 9012, 62%) were not registered as a case in a jurisdictional ARF/RHD register. The majority of these cases were identified in the hospital data, with additional data sources only contributing to a marginal number of

Table 2 Cumulative Frequencies Of Persons Included In The Expanded Cohort By Data Source

	All Ages		Under 60	
Register	5281	(12.6%)	5062	(35.7%)
Hospital	36,314	(86.3%)	8858	(62.4%)
Surgery	170	(0.4%)	35	(0.2%)
Other validation	135	(0.3%)	87	(0.6%)
Primary health care	94	(0.2%)	82	(0.6%)
Emergency department	70	(0.2%)	64	(0.5%)
Total	42,064	(100%)	14,188	(100%)

Note: Each person was uniquely counted towards a data source starting with ARF/RHD registers, hospital, etc.

additional cases. This also confirms that the analysis cohort (ARF/RHD registers, hospital and surgery data) provides a high level of case capture while maintaining a high level of comparability between jurisdictions.

Diagnosis dates for ARF were primarily retrieved from the ARF/RHD register records while most RHD onset dates were based on hospital records (see Tables 3 and 4). Hospital admissions identify an even higher proportion of people when considering people of all ages, especially for RHD reflecting the focus of ARF/RHD registers on monitoring younger patient cohorts. More detailed analysis will be undertaken by the ERASE Project team to compare case capture through register and hospital records.

Table 3 Cumulative Frequencies Of First-Ever Acute Rheumatic Fever (ARF) Diagnosis Dates In The Expanded Cohort, By Data Source

	All Ages		Under 60	
Register	2882	(60.7%)	2877	(63.6%)
Hospital	4522	(95.2%)	4311	(95.2%)
Other	4748	(100.0%)	4526	(100.0%)
No ARF record	37,316	(88.7%)	9662	(68.1%)
Total	42,064	(100.0%)	14,188	(100.0%)

Note: Each ARF episode was uniquely counted towards a data source starting with ARF/RHD registers, hospital, etc.

Table 4 Cumulative Frequencies Of Rheumatic Heart Disease (RHD) Onset Dates In The Expanded Cohort, By Data Source

	All Ages		Under 60	
Register	2858	(7.2%)	2691	(22.4%)
Hospital	38,797	(97.7%)	11,323	(94.4%)
Other	39,698	(100.0%)	12,000	(100.0%)
No RHD record	2366	(5.6%)	2188	(15.4%)
Total	42,064	(100.0%)	14,188	(100.0%)

Note: Each RHD onset diagnosis was uniquely counted towards a data source starting with ARF/RHD registers, hospital, etc.

Cohort Profile

Figure 2 shows the age distribution of first-ever ARF and RHD diagnoses by Indigenous status and diagnosis type for the analysis cohort. This figure clearly indicates the presence of two distinct affected populations: a sizable number of older non-Indigenous historical cases of RHD and a nominally smaller, but proportionally larger, population of much younger Indigenous persons for whom the available data can often trace the disease progression from ARF to RHD. While identifying both patient populations is important, in terms of prevention of disease progression and prospective case management, the younger cohort is of particular policy relevance. Therefore, much of the analyses conducted as part of the ERASE Project will focus on persons under 60 years at the time of initial diagnosis. Table 5 gives an overview of the descriptive characteristics of this age group for the three defined study cohorts.

The analysis cohort was more than double the size of the register-only cohort (Table 5; using all age groups rather than only under-60 years makes this increase even larger, Figure 1). As suggested in the previous section, there was only a small increase in case identification when moving from the analysis cohort to the expanded cohort. The additional cases in the analysis and expanded cohorts were primarily RHD-only cases. We found that Indigenous people were over-represented in the register cohort, even for younger age groups. Furthermore, in all three cohorts, Indigenous people were the majority of cases. The analysis and expanded cohorts also had a larger proportion of individuals aged over 45 years. The sex distribution was consistent across cohorts, with almost two-thirds of cases being female.

The geographical distribution of ARF and RHD cases varied substantially by Indigenous status (see Figure 3). A north–south gradient was evident for Indigenous people with ARF or RHD who typically reside in the more remote north of Australia. Non-Indigenous people with ARF or RHD, on the other hand, were concentrated largely in metropolitan areas.

Discussion

The ERASE Project originated out of the need for comprehensive, national ARF/RHD data by the END RHD Centre of Research Excellence (endrhd.telethon.kids.org.au, END RHD CRE) and END RHD (endhrd.org). The project has assembled a large and rich linked dataset that facilitates studying the epidemiology of ARF and RHD in Australia

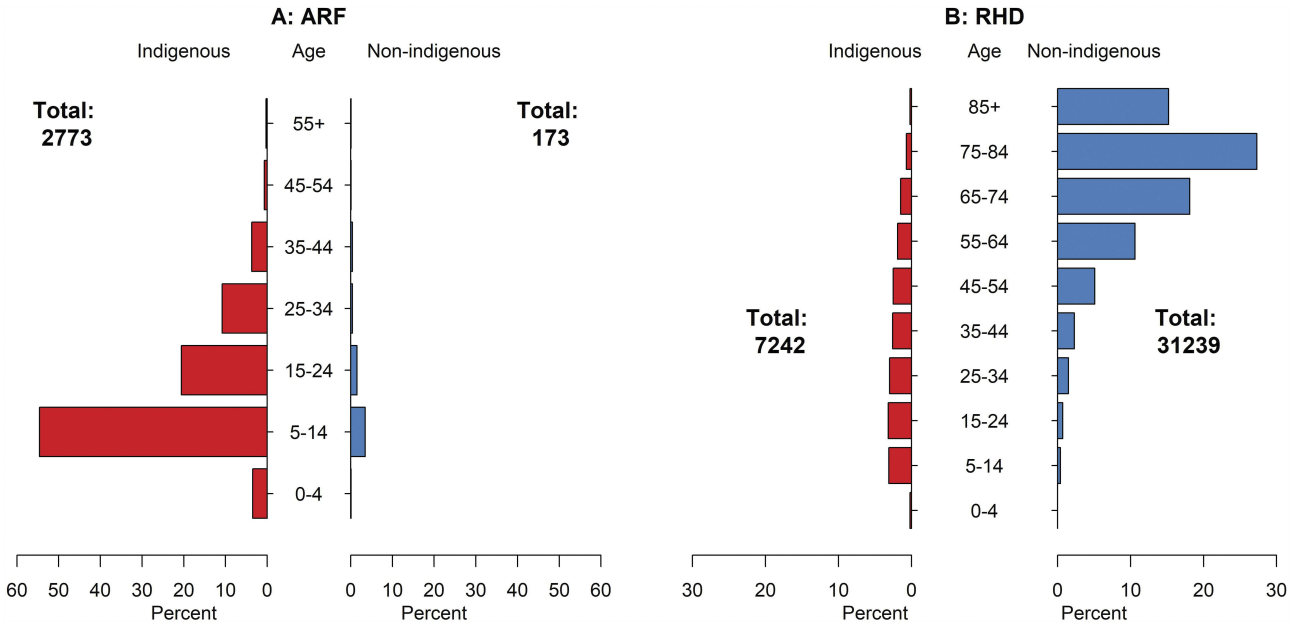


Figure 2 Age distribution of cases at time of initial (**A**). Acute rheumatic fever (ARF) and (**B**). Rheumatic heart disease (RHD) diagnoses, by Indigenous status (2001–2017, mid-year).

with unprecedented detail and rigor. The data include information from ARF/RHD registers, hospital data, death records, detailed surgical data, and other sources. In addition to collating a rich data source, methodological work is a key component of the ERASE Project to improve upon existing estimates of the burden of ARF/RHD. The proposed case and cohort identification provide the basis for robust and reliable epidemiological analyses of ARF and RHD.

This paper presents the first realistic estimates of the scale of the ARF/RHD patient population in Australia. We present estimates of the gap in case identification by the ARF/RHD registers. Including records for non-registered ARF/RHD patients results in more non-Indigenous and older patients being represented. However, the largest disease burden remains with young, Indigenous Australians.

There is often an underestimated level of social and emotional burden carried by families and communities that is hidden in routine epidemiological and quantitative reports. Acknowledging this gap, we encourage considering any quantitative estimates alongside qualitative material from associated studies that reflects the lived experience of these conditions.[49–51] More technical limitations of our study include those generally applicable to administrative data, including reliance on its availability, properties of the data reflecting its collection for purposes other than research, and challenges to consistent and accurate data entry. Because of data aggregation and limited data availability, we are unable to reliably further discern differences among population categories. With regard to RHD in particular, we have developed a prediction model[42] to select cases from a pool of RHD-coded cases. Nonetheless, we face the well-known challenge of only observing point-in-time assessments of a person's disease status while the chronic nature of the disease is incremental and unobserved. We lack access to primary care data for most of the study population. Tracking of patients across jurisdictions is only possible for NT and SA; therefore, the lack of cross-jurisdictional linkage across all data collections may result in double counting of some cases and events could be missed during follow-up if they occurred in another jurisdiction. We also acknowledge that the clinical complexity of ARF/RHD diagnosis and management poses particular challenges for deriving reliable descriptive and inferential estimates on its epidemiology and etiology.

Appropriate Indigenous input and oversight at every stage of the project cycle has been an integral focus for the ERASE Project. The Indigenous Advisory group linked to the END RHD CRE has provided guidance, support, and input into the methods, where needed. The Indigenous investigators on the team (DB, KG, VW) and the project lead (JK) have jointly developed a publication policy to operationalize how contributions of Indigenous researchers and stakeholders are incorporated into authorship and publications arising from the project. An important aspect is incorporating Aboriginal and Torres Strait Islander understandings and values in the publication processes. In the future, in support of the growing recognition of the rights of Indigenous peoples to sovereignty

Table 5 Descriptive Profile Of Study Cohorts For Patients Under 60 Years At The Time Of First ARF Or RHD Diagnosis, N (%)

	Register Cohort (n=5049)		Analysis Cohort (n=12907)		Expanded Cohort (n=14188)	
Population category						
Indigenous	4588	(91%)	7184	(56%)	7818	(55%)
ILICs	240	(5%)	1828	(14%)	1927	(14%)
Other Australian	219	(4%)	3821	(30%)	4367	(31%)
Sex						
Male	1951	(39%)	4695	(36%)	5374	(38%)
Female	3097	(61%)	8211	(64%)	8813	(62%)
Age at first ARF/RHD diagnosis: median (SD)	17	(14.11)	38	(17.20)	36	(17.50)
Age groups at first ARF/RHD diagnosis						
0-4	119	(2%)	154	(1%)	231	(2%)
5-14	2047	(41%)	2163	(17%)	2667	(19%)
15-24	1129	(22%)	1722	(13%)	1955	(14%)
25-45	1262	(25%)	3744	(29%)	4034	(28%)
45-59	492	(10%)	5124	(40%)	5301	(37%)
Diagnosis at end of follow-up						
ARF only	1423	(28%)	1275	(10%)	2188	(15%)
RHD only	2104	(42%)	9961	(77%)	9662	(68%)
Both ARF and RHD	1522	(30%)	1671	(13%)	2338	(16%)
Jurisdiction of residence at first ARF/RHD diagnosis						
NSW	51	(1%)	2128	(23%)	3235	(23%)
NT	2265	(45%)	3436	(27%)	3712	(26%)
QLD	1880	(37%)	3674	(28%)	4073	(29%)
SA	83	(2%)	596	(5%)	637	(4%)
WA	764	(15%)	1723	(13%)	1962	(14%)

Abbreviation: ILICs, Immigrants from low- and lower-middle-income countries.

over their health data, the ERASE project will seek to build statistical and research capacity in Indigenous communities particularly affected by RHD. The development of Indigenous-owned, determined and controlled research processes, data infrastructures and protocols (governance)[52] is essential to support Indigenous communities in advocating for better health and health care.

In addition to the core analysis team and the existing national and international network of researchers interested in ARF/RHD who are involved in an advisory or collaborative role, the ERASE Project is open to collaborations with other interested parties to conduct further analyses and combine and compare data on ARF/RHD globally. These are sought through the following channels:

- Interested researchers in Australia can approach the study team to analyze data for research questions complementary to the project's existing aims.
- Researchers with access to data on ARF/RHD, in particular internationally, are invited to collaborate on comparative analyses (in particular, international data on validated RHD cases for joint analyses with the study's Australian validated dataset and for external validation of some of the study's methodological work).
- National and international experts on ARF/RHD from all backgrounds and stakeholders, including patients, are invited to provide feedback on our analytical and methodological work on ARF/RHD burden estimation and related work.

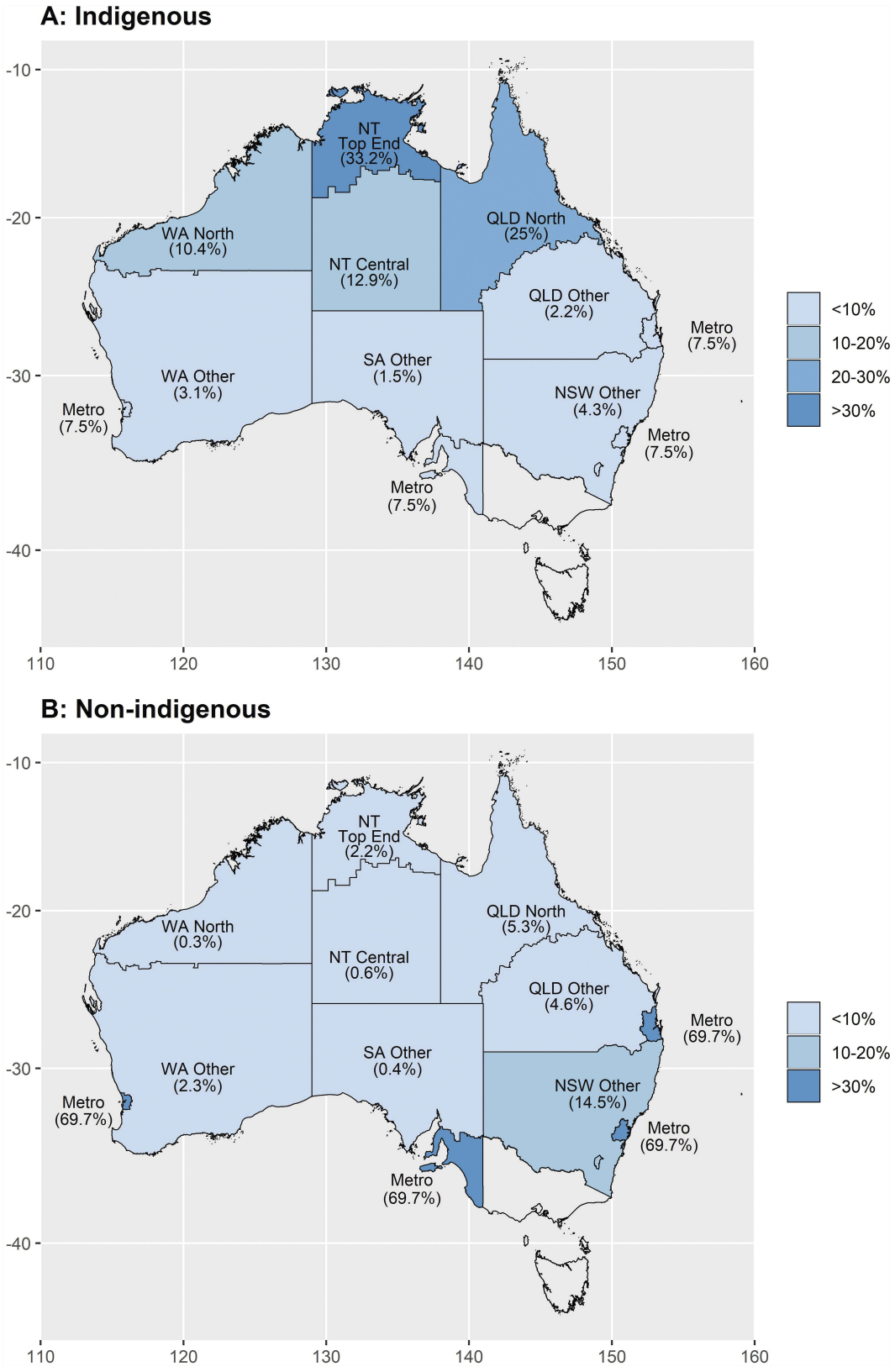


Figure 3 Geographical distribution of the Indigenous (**A**) and non-Indigenous (**B**) population diagnosed with either acute rheumatic fever or rheumatic heart disease at the time of the first diagnosis.

Conclusion

This work provides a detailed account of the compilation of the ERASE database incorporating the information available from different data sources so as to allow comparison thereof. Compiling the ERASE database was a critical and complex first step towards generating a reliable evidence base for studying ARF and RHD in Australia. It also provides an essential baseline for future disease monitoring at the quasi-national level. Future analyses under the ERASE Project will provide detailed morbidity and mortality estimates as well as analyses of disease progression, outcomes, adherence to prophylaxis, RHD-related surgery and economic costs, as well as their determinants.

We anticipate that the findings from the ERASE Project will contribute to ending ARF/RHD as a public health priority in Australia and reducing the global burden of the disease.

Ethics Approval

Human Research Ethics Committees of the Health Departments (and for NT: Menzies School of Health Research) of WA, SA, NT, QLD, and NSW provided approval for the ERASE Project. Aboriginal Ethics Committees from WA, SA, NT, and NSW also approved the study, after support letters from peak bodies of the Aboriginal Community Controlled Health Services. Supplementary Table 2 provides the list of committees and Indigenous organization that provided support.

Data Availability

The ERASE database cannot be shared publicly. Australian-based researchers can apply to the ERASE Project team with a proposal to analyze a pertinent research question using the ERASE Project data, subject to internal and ethics approval of the investigator and their research plans.

Acknowledgments

Other members of the ERASE Collaboration have been central to the feasibility and integrity of the project. These include Melanie Greenland and Ingrid Stacey who assisted with harmonization of the demographic data across datasets and jurisdictions; Professor Jonathan Carapetis, Prof Alex Brown and Prof Chris Reid who are associate investigators on ERASE and also investigators on the End RHD Centre for Research Excellence; Mellise Anderson who represents the RHD managers across the five jurisdictions; other investigators on the ERASE grant including Drs Angelita Martini, Kalinda Griffiths, Jess de Dassel, Fadwa Al-Yaman and Anne Russell; Sara Noonan and Catalina Lizama who have been contributed enormously to management of various components of the ERASE project; and Vicki Wade who provides significant cultural input into the research and dissemination process.

The authors value the support/endorsement provided to the project by the following peak bodies representing the Aboriginal Community Controlled Health sector: Aboriginal Medical Services Alliance Northern Territory, Kimberley Aboriginal Medical Service (regional Western Australian peak body) and Aboriginal Health Council of South Australia, Aboriginal Health and Medical Research Council (NSW). We also received support from the Aboriginal divisions of QLD and WA Health Departments. We are committed to providing feedback to the said organizations and ensuring that the findings are accessible and provide the evidence needed for a policy that can reduce the burden of ARF and RHD in Australia.

We acknowledge that figures and other statistics represent the loss of health and human life with profound impact and sadness for people, families, community, and culture. We hope that the 'numbers story' emanating from this project can augment the 'lived stories' that reflect the voices of people with RHD and their families, thus jointly contributing to evidence to erase suffering from ARF and RHD in Australia.

The authors also wish to thank the staff of the data linkage units of the State and Territory governments (WA, SA-NT, NSW, QLD) for the linkage of the data. We thank the State and Territory Registries of Births, Deaths and Marriages, the State and Territory Coroners, and the National Coronial Information System for enabling Cause of Death Unit Record File data to be used for this project.

Further, we thank the data custodians and data managers for the provision of the following data:

- Inpatient hospital data (5 states and territories)
- Emergency department data (5 states and territories)
- RHD registers (5 states and territories)
- ANZ Society of Cardiac & Thoracic Surgeons database (single data source from 5 states and territories)
- Royal Melbourne Children's Hospital Paediatric Cardiac Surgery database (single data source for RHD paediatric patients from SA and NT receiving surgery in Melbourne)
- Primary health care data from NT Department of Health

Supplementary Table 1 provides the list of data sources and linkage units. Judith M Katzenellenbogen and Daniela Bond-Smith are co-first authors for this study.

Disclosure

Dr Judith M Katzenellenbogen reports grants from National Health and Medical Research Council, National Heart Foundation Australia, and HeartKids, during the conduct of the study. Ms Rebecca Seth reports grants from National Health and Medical Research Council, during the conduct of the study. Professor Anna Ralph reports grants from National Health and Medical Research Council, during the conduct of the study. The authors report no other conflicts of interest in this work.

References

1. World Health Organization. *Seventy-First World Health Assembly A71/25. Resolution Adopted on 25 May 2018 for Provisional Agenda 12.8 By Director General on 12 April 2018*. Geneva: World Health Organization; 2018.
2. Carapetis JR, Beaton A, Cunningham MW, et al. Acute rheumatic fever and rheumatic heart disease. *Nat Rev Dis Primers*. 2016;2:15084. doi:10.1038/nrdp.2015.84
3. Noonan S, Zurynski YA, Currie BJ, et al. A national prospective surveillance study of acute rheumatic fever in Australian children. *Pediatr Infect Dis J*. 2013;32(1):e26–e32. doi:10.1097/INF.0b013e31826faeb3
4. Lawrence JG, Carapetis JR, Griffiths K, Edwards K, Condon JR. Acute rheumatic fever and rheumatic heart disease: incidence and progression in the Northern Territory of Australia, 1997 to 2010. *Circulation*. 2013;128(5):492–501. doi:10.1161/CIRCULATIONAHA.113.001477
5. Carapetis JR, Steer AC, Mulholland EK, Weber M. The global burden of group A streptococcal diseases. *Lancet Infect Dis*. 2005;5(11):685–694. doi:10.1016/S1473-3099(05)70267-X
6. Seckeler MD, Hoke TR. The worldwide epidemiology of acute rheumatic fever and rheumatic heart disease. *Clin Epidemiol*. 2011;3:67–84. doi:10.2147/CLEP.S12977
7. Steer AC, Carapetis JR. Prevention and treatment of rheumatic heart disease in the developing world. *Nat Rev Cardiol*. 2009;6(11):689–698. doi:10.1038/nrcardio.2009.162
8. Watkins DA, Johnson CO, Colquhoun SM, et al. Global, regional and national burden of rheumatic heart disease, 1990-2015. *N Engl J Med*. 2017;377(8):713–722. doi:10.1056/NEJMoa1603693
9. Coffey PM, Ralph AP, Krause VL. The role of social determinants of health in the risk and prevention of group A streptococcal infection, acute rheumatic fever and rheumatic heart disease: a systematic review. *PLoS Negl Trop Dis*. 2018;12:e0006577. doi:10.1371/journal.pntd
10. Carapetis JR. Rheumatic heart disease in developing countries. *N Engl J Med*. 2007;357(5):439–441. doi:10.1056/NEJMp078039
11. Carapetis JR, Wolff DR, Currie BJ. Acute rheumatic fever and rheumatic heart disease in the Top End of Australia's Northern Territory. *Med J Austr*. 1996;164(3):146–149.
12. Colquhoun SM, Condon JR, Steer AC, Li SQ, Guthridge S, Carapetis JR. Disparity in mortality from rheumatic heart disease in indigenous Australians. *J Am Heart Assoc*. 2015;4(7). doi:10.1161/JAHA.114.001282
13. Webb R, Wilson N. Rheumatic fever in New Zealand. *J Paediatr Child Health*. 2013;49(3):179–184. doi:10.1111/j.1440-1754.2011.02218.x
14. Australian Bureau of Statistics. Estimates of aboriginal and Torres Strait Islander Australians, June 2016. Canberra: ABS; 2018. Contract No.: Table 3238. 0.55.001.
15. Australian Bureau of Statistics. Life tables for aboriginal and Torres Strait Islander Australians, 2015-2017. Canberra: ABS; 2019. Contract No.: Table 3302. 0.55.003.
16. Australian Institute of Health and Welfare. *Australian Burden of Disease Study: Impact and Causes of Illness and Death in Aboriginal and Torres Strait Islander People 2011*. Canberra: AIHW; 2016.
17. Aboriginal and Torres Strait Islander Social Justice Commissioner. Social justice report 2010. Sydney: Australian Human Rights Commission; 2010. Report No. 1/2011.
18. Griffiths K, Coleman C, Lee V, Madden R. How colonisation determines social justice and Indigenous health - a review of the literature. *J Popul Res*. 2016;33(1):9–30. doi:10.1007/s12546-016-9164-1
19. Marmot M. Social determinants and the health of Indigenous Australians. *Med J Aust*. 2011;194(10):512–513.
20. Analysis HP. *Evaluation of the Commonwealth Rheumatic Fever Strategy – Final Report*. Canberra: Primary Healthcare Branch, Commonwealth Department of Health; 2017.
21. Cannon J, Bessarab DC, Wyber R, Katzenellenbogen JM. Public health and economic perspectives on acute rheumatic fever and rheumatic heart disease: can we afford 'business as usual. *Med J Aust*. 2019;211(6). doi:10.5694/mja2.50318
22. Wyber R, Katzenellenbogen JM, Pearson G, Gannon M. The rationale for action to end new cases of rheumatic heart disease in Australia. *Med J Aust*. 2017;207(8):322–323. doi:10.5694/mja17.00246
23. Katzenellenbogen JM, Nedkoff L, Canon J, et al. Low positive predictive value of ICD-10 codes in relation to rheumatic heart disease: a challenge for global surveillance. *Int Med J*. 2019;49(3):400–403. doi:10.1111/imj.14221
24. Organization WH. *Rheumatic Fever and Rheumatic Heart Disease. Seventy-First World Health Assembly A71/25. Resolution Adopted on 25 May 2018 for Provisional Agenda 12.8 By Director General on 12 April 2018*. Geneva: WHO; 2018.
25. Frances M, Morphy H. Anthropological theory and government policy in Australia's Northern Territory: the hegemony of the "mainstream". *Am Anthropol*. 2013;115(2):174–187. doi:10.1111/aman.12002
26. Memmott P, Long S, Bell M, Taylor J, Brown D Between places: indigenous mobility in remote and rural Australia. Brisbane: Australian Housing and Urban Research Institute, Queensland Research Centre; 2004. Report No.: AHURI Position Paper No 81.
27. Spilsbury K, Rosman D, Alan J, Boyd JH, Ferrante AM, Semmens JB. Cross-border hospital use: analysis using data linkage across four Australian states. *Med J Aust*. 2015;202(11):582–586. doi:10.5694/mja14.01414
28. Andrews NE, Sundararajan V, Thrift AG, et al. Addressing the challenges of cross-jurisdictional data linkage between a national clinical quality registry and government held state and national health data. *ANZ J Pub Hlth*. 2016;40(5):436–442.
29. Data linkage expert advisory group. *A Review of Western Australia's Data Linkage Capabilities - Commissioned Report*. Perth: Government of Western Australia; 2016.
30. Australian and New Zealand Society of Cardiac and Thoracic Surgeons (ANZSCTS). National Cardiac Surgery Database. Melbourne: CCRE, Monash University; 2010. Available from: https://anzscts.org/national-database/. Accessed October 09, 2019.
31. Rheumatic Fever Working Party of the Medical Research Council of Great Britain and the Subcommittee of Principal Investigators of the American Council on Rheumatic Fever and Congenital Heart Disease, American Heart Association.. The evolution of rheumatic heart disease in children: five-year report of a co-operative clinical trial of ACTH, cortisone and aspirin. *Can Med Assoc J*. 1960;83(15):781–789.
32. National Heart Foundation of Australia and the Cardiac Society of Australia and New Zealand. *The Australian Guideline for Prevention, Diagnosis and Management of Acute Rheumatic Fever and Rheumatic Heart Disease*. 2nd ed. 2012.

33. Australian Health Ministers' Advisory Council. *Aboriginal and Torres Strait Islander Health Performance Framework 2012 Report*. Canberra: AHMAC; 2012.
34. Australian Institute of Health and Welfare. *Rheumatic Heart Disease and Acute Rheumatic Fever in Australia: 1996-2012*. Canberra: AIHW; 2013.
35. Cannon J, Roberts K, Milne C, Carapetis JR. Rheumatic heart disease severity, progression and outcomes: a multi-state model. *J Am Heart Assoc*. 2017;6(3). doi:10.1161/JAHA.116.003498
36. de Dassel JL, de Klerk N, Carapetis JR, Ralph AP. How many doses make a difference? An analysis of secondary prevention of rheumatic fever and rheumatic heart disease. *J Am Heart Assoc*. 2018;7(24): e010223. doi:10.1161/JAHA.118.008528
37. de Dassel JL, Fittock MT, Wilks SC, Poole JE, Carapetis JR, Ralph AP. Adherence to secondary prophylaxis for rheumatic heart disease is underestimated by register data. *PLoS One*. 2017;12(5):e0178264. doi:10.1371/journal.pone.0178264
38. He VY, Condon JR, Ralph AP, et al. Long-term outcomes from acute rheumatic fever and rheumatic heart disease: a data-linkage and survival analysis approach. *Circulation*. 2016;134(3):222–232. doi:10.1161/CIRCULATIONAHA.115.020966
39. Roberts K, Maguire G, Brown A, et al. Echocardiographic screening for rheumatic heart disease in high and low risk Australian children. *Circulation*. 2014;129(19):1953–1961. doi:10.1161/CIRCULATIONAHA.113.003495
40. Australian Consortium for Classification Development. *The International Statistical Classification of Diseases and Related Health Problems, Tenth Revision, Australian Modification (ICD-10-AM/ACHI/ACS)*. 10th ed. Darlinghurst (NSW): Independent Hospital Pricing Authority; 2017.
41. Fitz-Gerald JA, Ongzalima CO, Ng A, Greenland M, Sanfilippo FM, Hung J, Katzenellenbogen JM. A validation study: how predictive is a diagnostic coding algorithm at identifying rheumatic heart disease in Western Australian hospital data? *Heart Lung Circ*. In press 2019.
42. Bond-Smith D, Knuiman MW, de Klerk N, Nedkoff L, Anderson M, Katzenellenbogen J, editors. *A Mixed Effects Prediction Model for Ascertaining Rheumatic Heart Disease Status Based on Linked Hospital Records*. Perth: Australasian Epidemiology Association Scientific Meeting, AEA; 2018.
43. Australian Institute of Health and Welfare. Indigenous identification in hospital separations data– quality report. Canberra: AIHW; 2013. Report No.: Cat. no. IHW 90.
44. Australian Institute of Health and Welfare. Report on the use of linked data relating to aboriginal and Torres Strait Islander people. Canberra: AIHW; 2013. Report No.: Cat. no. IHW 92.
45. Thompson SC, Woods JA, Katzenellenbogen JM. The quality of Indigenous identification in administrative health data in Australia: insights from studies using data linkage. *BMC Med Inform Decis Mak*. 2012;12:133. doi:10.1186/1472-6947-12-114
46. Christensen D, Davis G, Draper G, Mitrou F, McKeown S, Lawrence D. Evidence for the use of an algorithm in resolving inconsistent and missing Indigenous status in administrative data collections. *Aust J Soc Issues*. 2014;49(4):423–443. doi:10.1002/j.1839-4655.2014.tb00322.x
47. World Bank. World Bank Country and Lending Groups. Washington (DC): World Bank. Available from: https://datahelpdesk.worldbank.org/knowledgebase/articles/906519-world-bank-country-and-lending-groups. Accessed October 09, 2019.
48. Australian Bureau of Statistics. Australian Statistical Geography Standard (ASGS). Canberra: ABS; 2011. Report No.: Cat Number 1216.0.
49. Belton S, Kruske SJPL, Sherwood J, et al. Rheumatic heart disease in pregnancy: how can health services adapt to the needs of Indigenous women? A qualitative study. *Aust N Z J Obstet Gynaecol*. 2018;58(4):425–431. doi:10.1111/ajo.12744
50. Haynes E, Marawili M, Marika BM, et al. Community-based participatory action research on rheumatic heart disease in an Australian Aboriginal homeland: evaluation of the 'On track watch' project. *Eval Program Plann*. 2019;74:38–53. doi:10.1016/j.evalprogplan.2019.02.010
51. Mincham CM, Toussaint S, Mak DB, Plant AJ. Patient views on the management of rheumatic fever and rheumatic heart disease in the Kimberley, a qualitative survey. *Austr J Rural Hlth*. 2003;11(6):260–265. doi:10.1111/j.1440-1584.2003.00531.x
52. Lovett R, Lee V, Kukutai T, Cormach D, Rainie SC, Walker J. Good data practices for Indigenous data sovereignty and governance. In: Daly A, Devitt SK, Mann M, editors. *Good Data*. Amsterdam: Institute of Network Cultures; 2019:26–36.

Clinical Epidemiology Dovepress

Publish your work in this journal

Clinical Epidemiology is an international, peer-reviewed, open access, online journal focusing on disease and drug epidemiology, identification of risk factors and screening procedures to develop optimal preventative initiatives and programs. Specific topics include: diagnosis, prognosis, treatment, screening, prevention, risk factor modification, systematic reviews, risk & safety of medical interventions, epidemiology & biostatistical methods, and evaluation of guidelines, translational medicine, health policies & economic evaluations. The manuscript management system is completely online and includes a very quick and fair peer-review system, which is all easy to use.

Submit your manuscript here: https://www.dovepress.com/clinical-epidemiology-journal