



Challenges and opportunities for using population health data to investigate cancer survivors' quality of life in Australia

Imogen Ramsey¹ · Nadia Corsini¹ · Amanda Hutchinson² · Julie Marker³ · Marion Eckert¹

Accepted: 21 February 2022 / Published online: 4 March 2022
© The Author(s) 2022

Abstract

There is a recognised need for reported national data that inform health policy, health professions, and consumers about the wellbeing of Australians with cancer and other chronic conditions. International initiatives have demonstrated the viability and benefits of utilising population-based cancer registries to monitor the prevalence and trajectory of health-related quality of life (HRQOL) outcomes among people with cancer. Establishing a similar level of monitoring in Australia would require timely access to health data collected by publicly funded, population-based cancer registries, and the capacity to link this information across jurisdictions. Combining information from different sources via data linkage is an efficient and cost-effective way to maximise how data are used to inform population health and policy development. However, linking health datasets has historically been highly restricted, resource-intensive, and costly in Australia due to complex and outdated legislative requirements, duplicative approval processes, and differing policy frameworks in each state and territory. This has resulted in significant research waste due to underutilisation of existing data, duplication of research efforts and resources, and data not being translated into decision-making. Recognising these challenges, from 2015 to 2017 the Productivity Commission investigated options for improving data availability and use in Australia, considering factors such as privacy, security, and intellectual property. The inquiry report recommended significant reforms for Australian legislation, including the creation of a data sharing and release structure to improve access to data for research and policy development purposes. This paper discusses (1) opportunities in HRQOL research enabled by data linkage, (2) barriers to data access and use in Australia and the implications for waste in HRQOL research, and (3) proposed legislative reforms for improving data availability and use in Australia.

Keywords Quality of life · Population health · Cancer · Surveillance

Background

Globally, population-based health data are being utilised to advance our understanding of trends in population health and generate timely insights for clinical practice and policy. However, in Australia, the use of population-based

surveillance datasets for research has been hampered by the complexity and duplication of state and national legislation and regulations regarding data access and linkage. As a result, we have not optimised the use or social benefits of health data for improving the health-related quality of life (HRQOL) of people affected by cancer or other chronic diseases.

Waste has been defined as “non-value” and “all efforts that do not add value” [1]. In its 2014 series “Increasing value, reducing waste”, The Lancet examined sources of avoidable waste in biomedical research, which were first highlighted by Chalmers and Glasziou in 2009 [2]. These sources of waste included studies that overlook important research questions [3]; that are poorly designed or conducted [4]; that are inefficiently regulated and managed [5]; that produce inaccessible information [6]; and that are not

✉ Imogen Ramsey
imogen.ramsey@unisa.edu.au

¹ Rosemary Bryant AO Research Centre, UniSA Clinical & Health Sciences, University of South Australia, City East Campus, GPO Box 2471, Adelaide, SA, Australia

² UniSA Justice & Society, University of South Australia, Adelaide, SA, Australia

³ Cancer Voices South Australia, Adelaide, SA, Australia

appropriately reported, disseminated, or translated into decision-making [7]. This paper focuses on inefficient research regulation and management practices as sources of various types of research waste. In addition to consuming valuable research resources, these processes prevent data from being accessed in a timely manner or at all, and therefore from being used to create value by advancing research or informing decision-making. Transparency and access to data was one of six quality dimensions agreed upon by international stakeholders for inclusion in the INcreasing QUality In patient-oriented academic clinical REsearch (INQUIRE) framework, which was developed to assist researchers in navigating the challenges of reducing waste and increasing value [8]. Ensuring that regulation and management are proportionate to risks is one of five priorities in The Lancet's REduce Waste And Reward Diligence (REWARD) campaign, which invites funders, regulators, organisations, publishers, researchers and other stakeholders to commit to reducing waste and maximising efficiency [9].

There is a recognised need for population-based data on the long-term HRQOL experienced by people with cancer [10]. This growing population is at risk of chronic health problems; comorbidity; adverse physical, psychological, and social outcomes; secondary cancers; and long-term and late effects of treatment, all of which can significantly impact on HRQOL [11, 12]. HRQOL is inversely associated with behavioural and lifestyle risk factors and frequently predicts morbidity, mortality, and healthcare utilisation more accurately than objective health measures [13]. Therefore, its surveillance at a population level can support disease burden prediction, resource allocation, and identification of public health issues [13]. Gaining an understanding of the prevalence and trajectory of HRQOL outcomes at a population level, and how they differ based on sociodemographic and clinical factors, can inform clinical practice, health policy, the design of appropriate health interventions and services, and the allocation of health resources to areas of highest priority [14].

In Australia, population-level data on cancer incidence, mortality, and survival are collected by publicly funded, population-based cancer registries, allowing for national reporting of these outcomes over time [10]. However, these indicators represent only a fraction of the impact of cancer on population health. We need an understanding of not just length of survival, but the long-term HRQOL experienced by people with cancer, to design health services and policy that meet their needs. Collecting HRQOL data using patient-reported outcomes (PROs; self-reports about a patient's health provided directly by the patient) at a population level presents a challenge, as this is typically beyond the remit or mandate of cancer registries [15]. In addition, legislative barriers to timely data linkage have inhibited the sharing of clinical registry data that could inform population-based

HRQOL research in Australia, stifling innovation in this area. This paper discusses (1) opportunities in HRQOL research enabled by data linkage, using an international example; (2) barriers to data access and use in Australia and the implications for waste in HRQOL research; and (3) proposed legislative reforms.

Opportunities in HRQOL research enabled by data linkage

Data linkage is the process of uniting data pertaining to an individual, group, location, or event, from multiple independent data sources, in a way that maintains privacy and confidentiality [16]. Combining data from different sources is an efficient and cost-effective way to maximise the use of existing data for informing population health and policy development, thereby reducing research burden and duplication of effort [16]. Additionally, data linkage can make detailed longitudinal patient data available and provide the infrastructure to investigate research questions requiring the use of multiple datasets [16, 17]. In the context of HRQOL research in cancer, linkage with datasets held by various levels of government can be used to examine associations of HRQOL with cancer incidence and mortality data, clinical data, and sociodemographic characteristics.

Internationally, a small number of initiatives have demonstrated the viability and benefits of collecting PROs from people with cancer utilising population-based cancer registries and linking this information with clinical and sociodemographic data, to understand the impact of cancer on HRQOL at a national scale [18]. In a systematic review, we identified seven international registries that had been established for ongoing surveillance of HRQOL among defined populations of cancer survivors. The review findings highlighted the viability of using population-based cancer registries to collect PRO data from cancer survivors. The data generated by these systems have been used for a variety of purposes, including informing disease management, clinical decision-making, health care planning, quality improvement, and understanding the impacts of cancer on population health [18].

A unique and well-established example of HRQOL surveillance is the Patient-Reported Outcomes Following Initial treatment and Long-term Evaluation of Survivorship (PROFILES) registry [19], which has collected comprehensive, longitudinal PRO data about physical, psychological, and social aspects of HRQOL from cancer survivors in the Netherlands since 2009 [18]. People diagnosed with cancer are invited by their treating physician to complete a voluntary questionnaire annually, for up to five years from primary treatment completion, allowing outcomes to be monitored over time [19]. Through data linkage with the Netherlands

Cancer Registry, which records clinical information about all cancers diagnosed in the Netherlands, PROs are linked with, and interpreted in relation to, participants' clinical and sociodemographic characteristics [19]. The comprehensive datasets generated by the PROFILES registry have enabled trends in HRQOL to be monitored longitudinally at a population level, and according to factors such as tumour type and stage, age, treatment received, and comorbidity. Additionally, the availability of corresponding PRO data from a normative non-cancer population allows comparison of HRQOL between population-based cohorts of cancer survivors and demographically matched controls [19].

Research studies analysing data from the PROFILES registry have generated invaluable insight into the course of HRQOL experienced by cancer survivors in the Netherlands, documented in over 100 research studies that explored and compared outcomes among different subgroups, and for different interventions. For example, these studies demonstrated that cancer survivors reported persistently lower HRQOL and functioning [20–22], and higher prevalence of fatigue [23–26], anxiety, and depressive symptoms [27, 28] than individuals without cancer. Moreover, they have identified associations between outcomes including fatigue, depressive symptoms, and poor HRQOL with all-cause mortality [29–31]. Besides linkage with clinical cancer registry data, PROFILES data have been linked with clinical trials data to examine possible trial effects on the long-term HRQOL of cancer survivor cohorts [15], and with pharmaceutical data to determine the impacts of medications on long-term HRQOL [15]. In addition to enhancing understanding of the long-term impacts of cancer on HRQOL and opening new fields of research [15], HRQOL research studies utilising linked data have informed practice and policy for cancer survivors in the Netherlands [18]. For example, findings from the PROFILES registry have facilitated change in diverse areas including chemotherapy regimens, models of supportive care, information provision, and travel insurance [18].

Barriers to data access and use, and implications for waste in HRQOL research

Compared with practices for data sharing in the Netherlands, linkage and integration of datasets is inadequate in Australia, where barriers to accessing and linking data sources for research purposes have been well documented [16, 32–35]. For a combination of legal, technical, and institutional reasons, Australia has historically used health information poorly compared with other developed countries [36]. Factors that restrict Australia's sharing of data include complex legislative requirements, overly risk-averse approval

processes, the lack of a whole of government approach, and differing policy frameworks in each jurisdiction [37]. The combination of these factors has resulted in research waste due to lengthy and complex processes that consume valuable research resources, duplication of research efforts, and research data not being used to address important research questions or translated into decision-making to benefit the public.

To improve data linkage and sharing within Australia, the Population Health Research Network (PHRN) was established in 2009 [34]. This network comprises national data linkage units in all Australian jurisdictions, which have the capacity to provide linked data on a national scale [34, 38]. This investment in Australian data linkage infrastructure has improved access to linked data [39] and seen an increase in the number of research publications involving the use of linked health and human services data, from 72 publications in 2009–2010 to 199 in 2016–2017 [34]. Studies using PHRN linked data have influenced health policy and service provision in the areas of cancer, diabetes, heart and cardiovascular disease, and paediatric health [40]. The results from one study that used linked data, for example, prompted changes to clinical guidelines and education about cancer risk and dosage and exposure practices for computed tomography scans [41]. The contribution of PHRN-related data to reductions in cancer burden (i.e. projected disability-adjusted life years resulting from cancer) in Australia has been estimated to exceed 0.5% by 2034, associated with a net economic benefit of \$7.1 billion [42]. Despite growth in publications from research using linked data over the past 8 years, major barriers to the timely and effective use of data linkage in Australia remain [34, 38]. The jurisdictional data linkage facilities were established along with governance protocols to protect individual privacy and data confidentiality to the highest possible ethical standard [32]. However, approval processes that previously existed to protect privacy and confidentiality remained in place [32]. Researchers are therefore required to obtain multiple approvals including from data linkage centres, data custodians, and ethics committees [32]. Due to the complexity, duplication, and lack of cohesion of the procedures required to undertake data linkage, it remains a time consuming, costly, and resource-intensive process, particularly linkage of cross-jurisdictional datasets [32, 35]. For a national study, obtaining required approvals can require liaison with numerous data custodians, up to seven data linkage centres, and more than 10 ethics committees [32], with expected timelines exceeding three years [16].

As stated in Cancer Australia's National Cancer Data Strategy [10] "while due regard for ethical standards and privacy protocols is very important, the need for multiple clearances and multiple approval processes is a major source of inefficiency and often, a barrier to progress" [10].

Indeed, while the increasing number of data sources used for population health research necessitates due consideration of consumers' privacy and confidentiality [13], the time and resources expended in achieving data linkage undermine the potential value of research for population health [32]. The inefficiencies associated with data linkage in Australia are exacerbated by the legal requirement that datasets resulting from linked Commonwealth data (e.g. from the Pharmaceutical Benefits Program or Medicare Benefits Scheme) be destroyed once a specific project is completed [37]. The waste inherently associated with the regular destruction and recreation of datasets that are linked at a high cost is significant.

Despite a decade of investment in infrastructure and funding, Australian health data are still an underutilised resource due to poor data sharing integration practices [33]. A consequence of the barriers cited is that they create disincentives to invest in systems similar to the PROFILES registry that would allow HRQOL to be assessed at a population level, to answer questions that underpin prudent cancer control strategies. These opportunity costs represent “unnoticed and unquantifiable waste, such that important research is identified but never addressed” [5]. Instead of investing in HRQOL data infrastructure to support multiple studies, there is currently a reliance on collecting data de novo for discrete studies, resulting in potential duplication of resources across the research pipeline and studies being underpowered. Ultimately, these sources of waste limit our ability to answer questions about population health and translate research data into meaningful policy action [44]. In 2008, Cancer Australia called for the gap in HRQOL data to be addressed as part of its National Cancer Data Strategy [10] but, because of these barriers, we are no closer to bridging this gap in 2021.

Waste caused by inaccessible data and inefficiencies in regulatory processes are not unique to Australia, as *The Lancet's* series highlighted [5]. Internationally, governments and organisations are seeking to streamline processes and enable greater sharing of data across public and private sectors, to encourage innovation and improvement in services across various industries, including health [43]. In the United Kingdom (UK), 21 academic and health institutions formed a dedicated research institute, the Farr Institute (now Health Data Research UK), to facilitate the safe use of electronic health records and administrative data for research purposes. The Farr Institute facilitated the use of population-based datasets in 593 published papers between 2013 and 2018, a time of profound societal and policy change within the UK in relation to health data science [45]. In 2018, European Union (EU) Member states launched Information for Action (InfAct), a joint initiative to improve the availability of health data and strengthen health information activities across Europe [46]. An InfAct study published in 2020

reported that data linkage is performed widely across EU countries for health status monitoring, policy development, and scientific research [47]. However, the authors found that more flexible governance frameworks, specific mandates to ensure data availability, and consistent interpretation of the EU General Data Protection Regulations (GDPR) are required to support and enhance data linkage capabilities [47].

Legislative reforms: a solution?

Realising the potential for national data linkage could enable valuable insights into HRQOL outcomes and their associations with medical treatments, health services, and demographic characteristics; information that is currently not available in Australia. In response to the reported challenges of data linkage, Australian researchers have called for more streamlined application and approval processes, less duplication, greater collaboration between jurisdictions, and increased infrastructure and funding support [16, 32–34]. Because consumers have the most to gain from reductions in research waste and inefficiency, they too should be involved in decisions about the need for, and extent of, the effects of regulation and management on research [5]. Support for enabling better use of Australia's health data in research appears to be evident among consumers, who are willing to share data in exchange for outputs that are meaningful [48]. In a recent survey, 83% of participants indicated they were willing to share their de-identified health data to advance medical research [49], consistent with international evidence that most patients consent to the use of secondary data for record linkage [50]. Australian consumers have reported that they are more willing to share data when it is being used for individual or public good, and when they understand how it is being used [48]. However, consumers share differing views and understand the concept of data linkage to differing extents, highlighting the need for transparent and two-way communication about the processes, benefits, and risks involved [48]. A whole of government approach to using health data, that is co-designed with consumers, has been identified as key to increasing the consistency, quality, and value of data linkage and use in Australia [48].

Recognising the need to improve opportunities for data access and use in Australia while considering factors such as privacy, security, and intellectual property, from 2015 to 2017 the Productivity Commission conducted a broad investigation into the benefits and costs of options for improving data availability and use [37]. The inquiry report recommended significant changes for Australia's open government agenda and the rights of consumers to data, to catch up with achievements in competing economies such as the UK, the US, New Zealand, Canada, France, and The Netherlands

[37]. A key component of the recommended reforms was the creation of a data sharing and release structure, to provide better and more timely access to data for research and policy development purposes [37]. The proposed new legislation for data sharing and release has been developed by the Office of the National Data Commissioner, through extensive stakeholder consultation [51]. The proposed laws will overwrite some 500 data secrecy and confidentiality provisions in 175 pieces of existing legislation [52]. However, data will only be released if it is related to improving policy, program evaluation, service delivery and research and development [52].

How the proposed laws will translate into policy to inform health research, and whether they will help to reduce waste and increase value and opportunities for innovation in HRQOL research remains to be seen. As the Productivity Commission's 2017 inquiry report found, a risk-averse culture may pose a greater threat to data sharing and access than legislation itself [37]. In the past, uncertainty among data collectors and custodians about how to interpret legislation such as the Privacy Act has led to overly cautious approaches to data management and complex and inefficient approval processes designed to meet legislative requirements [37]. Avoiding the same pitfalls in the reform of Australia's data infrastructure will therefore require not only a simplified legislative framework, but a substantial shift away from an entrenched culture of risk aversion [37].

In 2018, Ford and colleagues reported similar challenges in the UK, following six years of sustained investment in improving use of health data [53]. Obstacles to data access were cited as one of the key barriers to realising an inclusive, transparent, and innovative vision for health data science [53]. Specifically, the authors reported that data access was hampered by inconsistent use of national standards, despite the introduction of legislation designed to facilitate data science, such as the GDPR and the UK Digital Economy Act 2017 [53]. Similar to the Productivity Commission's Australian findings, Ford and colleagues linked these barriers to a broader culture of risk aversion in the UK, perpetuated by inconsistencies in how legislation is interpreted and applied [53].

In February 2021, the UK government announced a review into the efficient and safe use of health data for research to complement a proposed data strategy for health and social care. The draft strategy (*Data saves lives: reshaping health and social care with data (draft)*), released in June 2021, outlines the Secretary of State's vision for using data to improve population health [54]. Proposed legislative changes are among an extensive list of commitments made in the report, with others including improving information governance, developing guidance and frameworks for data sharing and transparency, enabling members of the public to access their own health data, providing education and training for staff, reducing data collection burden, building

analytical and data science capability, working with stakeholders and the public to implement legislative reform, modernising data architecture, and supporting innovation [54]. In outlining these commitments, the draft strategy proposes to support the creation of a strong data infrastructure and the implementation of not only legislative, but behavioural, cultural, organisational, and policy change.

Conclusions

Inefficient regulation and management of health data creates the types of waste that Al-Shahi Salman and colleagues argue arise “from questions being overlooked or unnecessarily addressed, research being underpowered or done too slowly, and research being too costly” [5]. Consistent with global trends, there is a national imperative to achieve efficient and timely health data linkage, to optimise how data are used to inform and improve the provision of health services in Australia. The projected increase in the number of people diagnosed with, and surviving, cancer will create greater future demand across health systems. In the context of this challenge, understanding the long-term impacts of cancer on HRQOL will become increasingly important to direct resources to where they can make the most difference. Australia's proposed legislative reforms are designed to maximise the value of public sector data for service delivery and research. It is hoped that the implementation of these reforms will reduce the fragmentation and duplication of research activities, enable more timely sharing of data for research, and expose existing data to a wider range of innovative ideas and approaches, to increase value and reduce waste. However, as the UK's new data strategy suggests, successful implementation will require the complex task of fostering a culture of data sharing and transparency. Data are wasted if they are collected and not used to influence decision-making. Increasing the availability and use of data will ultimately enable more research to add to international literature on population health, and inform evidence-based decision-making for clinical practice, service design, and health policy. Progress across these areas is critical for improving health systems and HRQOL outcomes for the global population affected by cancer and other chronic diseases.

Funding Open Access funding enabled and organized by CAUL and its Member Institutions. Not applicable.

Availability of data and material Not applicable.

Code availability Not applicable.

Declarations

Conflict of interest Not applicable.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

References

- Chroneer, D., & Wallström, P. (2016). Exploring waste and value in a lean context. *International Journal of Business and Management*, 11(10), 282–297.
- Chalmers, I., & Glasziou, P. (2009). Avoidable waste in the production and reporting of research evidence. *The Lancet*, 374(9683), 86–89.
- Chalmers, I., Bracken, M. B., Djulbegovic, B., Garattini, S., Grant, J., Gülmezoglu, A. M., et al. (2014). How to increase value and reduce waste when research priorities are set. *The Lancet*, 383(9912), 156–165.
- Ioannidis, J. P., Greenland, S., Hlatky, M. A., Khoury, M. J., Macleod, M. R., Moher, D., et al. (2014). Increasing value and reducing waste in research design, conduct, and analysis. *The Lancet*, 383(9912), 166–175.
- Salman, R.A.-S., Beller, E., Kagan, J., Hemminki, E., Phillips, R. S., Savulescu, J., et al. (2014). Increasing value and reducing waste in biomedical research regulation and management. *The Lancet*, 383(9912), 176–185.
- Chan, A.-W., Song, F., Vickers, A., Jefferson, T., Dickersin, K., Gøtzsche, P. C., et al. (2014). Increasing value and reducing waste: Addressing inaccessible research. *The Lancet*, 383(9913), 257–266.
- Glasziou, P., Altman, D. G., Bossuyt, P., Boutron, I., Clarke, M., Julious, S., et al. (2014). Reducing waste from incomplete or unusable reports of biomedical research. *The Lancet*, 383(9913), 267–276.
- von Niederhäusern, B., Guyatt, G. H., Briel, M., & Pauli-Magnus, C. (2018). Academic response to improving value and reducing waste: A comprehensive framework for INcreasing QUALity In patient-oriented academic clinical REsearch (INQUIRE). *PLoS Medicine*, 15(6), e1002580.
- Moher, D., Glasziou, P., Chalmers, I., Nasser, M., Bossuyt, P. M., Korevaar, D. A., et al. (2016). Increasing value and reducing waste in biomedical research: Who's listening? *The Lancet*, 387(10027), 1573–1586.
- Cancer Australia. (2008). A National Cancer Data Strategy for Australia: A collaborative approach to improving cancer outcomes through cancer data. Commonwealth of Australia.
- Miller, K. D., Nogueira, L., Mariotto, A. B., Rowland, J. H., Yabroff, K. R., Alfano, C. M., et al. (2019). Cancer treatment and survivorship statistics, 2019. *CA: A Cancer Journal for Clinicians*, 69(5), 363–385.
- Fowler H, Belot A, Ellis L, Maringe C, Luque-Fernandez MA, Njagi EN, et al. Comorbidity prevalence among cancer patients: a population-based cohort study of four cancers. *BMC Cancer*. 2020;20(2).
- Mamiya, H., Lix, L. M., Gardner, W., Bartlett, S. J., Ahmed, S., & Buckeridge, D. L. (2017). Montreal Accord on Patient-Reported Outcomes (PROs) use series—Paper 5: Patient-reported outcomes can be linked to epidemiologic measures to monitor populations and inform public health decisions. *Journal of Clinical Epidemiology*, 89, 142–147.
- Thong, M. S., Mols, F., Stein, K. D., Smith, T., Coebergh, J. W., & van de Poll-Franse, L. V. (2013). Population-based cancer registries for quality-of-life research: A work-in-progress resource for survivorship studies? *Cancer*, 119(S11), 2109–2123.
- Thong, M. S. Y., Mols, F., Doege, D., van de Poll-Franse, L., & Arndt, V. (2018). Population-based cancer survivorship research: Experiences from Germany and the Netherlands. *Journal of Cancer Policy*, 15, 87–91.
- Andrew, N. E., Sundararajan, V., Thrift, A. G., Kilkenny, M. F., Katzenellenbogen, J., Flack, F., et al. (2016). Addressing the challenges of cross-jurisdictional data linkage between a national clinical quality registry and government-held health data. *Australian and New Zealand Journal of Public Health*, 40(5), 436–442.
- Olver, I. N. (2014). Linking data to improve health outcomes. *Medical Journal of Australia*, 200(7), 368–369.
- Corsini, N., Fish, J., Ramsey, I., Sharplin, G., Flight, I., Damarell, R., et al. (2017). Cancer survivorship monitoring systems for the collection of patient-reported outcomes: A systematic narrative review of international approaches. *Journal of Cancer Survivorship*, 11(4), 486–497.
- van de Poll-Franse, L. V., Horevoorts, N., Eenbergen, M. V., Denollet, J., Roukema, J. A., Aaronson, N. K., et al. (2011). The Patient Reported Outcomes Following Initial treatment and Long term Evaluation of Survivorship registry: Scope, rationale and design of an infrastructure for the study of physical and psychosocial outcomes in cancer survivorship cohorts. *European Journal of Cancer*, 47(14), 2188–94.
- Husson, O., Haak, H. R., Buffart, L. M., Nieuwlaat, W.-A., Oranje, W. A., Mols, F., et al. (2013). Health-related quality of life and disease specific symptoms in long-term thyroid cancer survivors: A study from the population-based PROFILES registry. *Acta Oncologica*, 52(2), 249–258.
- Husson, O., Prins, J. B., Kaal, S. E., Oerlemans, S., Stevens, W. B., Zebrack, B., et al. (2017). Adolescent and young adult (AYA) lymphoma survivors report lower health-related quality of life compared to a normative population: Results from the PROFILES registry. *Acta Oncologica*, 56(2), 288–294.
- Mols, F., Oerlemans, S., Vos, A. H., Koster, A., Verelst, S., Sonneveld, P., et al. (2012). Health-related quality of life and disease-specific complaints among multiple myeloma patients up to 10 yr after diagnosis: Results from a population-based study using the PROFILES registry. *European Journal of Haematology*, 89(4), 311–319.
- Husson, O., Mols, F., van de Poll-Franse, L. V., & Thong, M. S. Y. (2015). The course of fatigue and its correlates in colorectal cancer survivors: A prospective cohort study of the PROFILES registry. *Supportive Care in Cancer*, 23(11), 3361–3371.
- Husson, O., Nieuwlaat, W.-A., Oranje, W. A., Haak, H. R., van de Poll-Franse, L. V., & Mols, F. (2013). Fatigue among short-and long-term thyroid cancer survivors: Results from the population-based PROFILES registry. *Thyroid*, 23(10), 1247–1255.
- Husson, O., Mols, F., van de Poll-Franse, L., de Vries, J., Schep, G., & Thong, M. S. (2015). Variation in fatigue among 6011 (long-term) cancer survivors and a normative population: A study from the population-based PROFILES registry. *Supportive Care in Cancer*, 23(7), 2165–2174.

26. Oerlemans, S., Mols, F., Issa, D. E., Pruijt, J., Peters, W. G., Lybeert, M., et al. (2013). A high level of fatigue among long-term survivors of non-Hodgkin's lymphoma: Results from the longitudinal population-based PROFILES registry in the south of the Netherlands. *Haematologica*, *98*(3), 479–486.
27. Mols, F., Schoormans, D., de Hingh, I., Oerlemans, S., & Husson, O. (2018). Symptoms of anxiety and depression among colorectal cancer survivors from the population-based, longitudinal PROFILES Registry: Prevalence, predictors, and impact on quality of life. *Cancer*, *124*(12), 2621–2628.
28. Oerlemans, S., Mols, F., Nijziel, M. R., Zijlstra, W. P., Coebergh, J. W. W., & van de Poll-Franse, L. V. (2014). The course of anxiety and depression for patients with Hodgkin's lymphoma or diffuse large B cell lymphoma: A longitudinal study of the PROFILES registry. *Journal of Cancer Survivorship*, *8*(4), 555–564.
29. Mols, F., Husson, O., Roukema, J.-A., & van de Poll-Franse, L. V. (2013). Depressive symptoms are a risk factor for all-cause mortality: Results from a prospective population-based study among 3,080 cancer survivors from the PROFILES registry. *Journal of Cancer Survivorship*, *7*(3), 484–492.
30. Husson, O., de Rooij, B. H., Kieffer, J., Oerlemans, S., Mols, F., Aaronson, N. K., et al. (2020). The EORTC QLQ-C30 Summary Score as prognostic factor for survival of patients with cancer in the “real-world”: Results from the population-based PROFILES registry. *The Oncologist*, *25*(4), e722–e732.
31. Adam, S., van de Poll-Franse, L. V., Mols, F., Ezendam, N. P., de Hingh, I. H., Arndt, V., et al. (2019). The association of cancer-related fatigue with all-cause mortality of colorectal and endometrial cancer survivors: Results from the population-based PROFILES registry. *Cancer Medicine*, *8*(6), 3227–3236.
32. Mitchell, R. J., Cameron, C. M., McClure, R. J., & Williamson, A. M. (2015). Data linkage capabilities in Australia: Practical issues identified by a Population Health Research Network ‘Proof of Concept project.’ *Australian and New Zealand Journal of Public Health*, *39*(4), 319–325.
33. Palamuthasingam, D., Johnson, D. W., Hawley, C., Pascoe, E., & Fahim, M. (2019). Health data linkage research in Australia remains challenging. *Internal Medicine Journal*, *49*(4), 539–544.
34. Young, A., & Flack, F. (2018). Recent trends in the use of linked data in Australia. *Australian Health Review*, *42*(5), 584–590.
35. Rowe, S. L., Stephens, N., Cowie, B. C., Nolan, T., Leder, K., & Cheng, A. C. (2019). Use of data linkage to improve communicable disease surveillance and control in Australia: Existing practices, barriers and enablers. *Australian and New Zealand Journal of Public Health*, *43*(1), 33–40.
36. OECD. (2015). OECD Reviews of Health Care Quality: Australia 2015: Raising Standards. OECD Publishing, Paris. <https://doi.org/10.1787/9789264233836-en>.
37. Commonwealth of Australia, Productivity Commission. (2017). Data availability and use. Canberra; 2017. Report No. 82. <https://www.pc.gov.au/inquiries/completed/data-access/report>.
38. Tew, M., Dalziel, K. M., Petrie, D. J., & Clarke, P. M. (2017). Growth of linked hospital data use in Australia: A systematic review. *Australian Health Review*, *41*(4), 394–400.
39. Flack, F., & Smith, M. (2019). Access to linked data: Challenges and progress. *Internal Medicine Journal*, *49*(10), 1339–1340.
40. Population Health Research Network. (2021). Impact stories 2021. <https://www.phrn.org.au/for-the-community/impact-stories/>.
41. Mathews, J. D., Forsythe, A. V., Brady, Z., Butler, M. W., Goergen, S. K., Byrnes, G. B., et al. (2013). Cancer risk in 680 000 people exposed to computed tomography scans in childhood or adolescence: Data linkage study of 11 million Australians. *BMJ*, *346*, 2360. <https://doi.org/10.1136/bmj.f2360>
42. Lateral Economics. Population Health Research Network (PHRN). (2017). Impact and return on investment. <https://www.phrn.org.au/media/81439/final-report-phrn-lateral-economics-oct-2017-final-clean.pdf>.
43. Phillips M, Molnár-Gábor F, Korbel JO, Thorogood A, Joly Y, Chalmers D, et al. Genomics: data sharing needs an international code of conduct. Nature Publishing Group; 2020.
44. Kruse, C. S., Goswamy, R., Raval, Y., & Marawi, S. (2016). Challenges and opportunities of big data in health care: A systematic review. *JMIR Medical Informatics*, *4*(4), e38.
45. Hemingway, H., Lyons, R., Li, Q., Buchan, I., Ainsworth, J., Pell, J., et al. (2020). A national initiative in data science for health: an evaluation of the UK Farr Institute. *International Journal of Population Data Science*, *5*(1), 112.
46. Abboud, L., Bogaert, P., Fehr, A., Urbanski, D., Tolonen, H., Noguer-Zambran, I., et al. (2018). The new Joint Action on Health Information: Information for action (InfAct)! *European Journal of Public Health*, *28*(4), 213.
47. Haneef, R., Delnord, M., Vernay, M., Bauchet, E., Gaidelyte, R., Van Oyen, H., et al. (2020). Innovative use of data sources: A cross-sectional study of data linkage and artificial intelligence practices across European countries. *Archives of Public Health*, *78*(1), 1–11.
48. Consumers Health Forum of Australia, NPS MedicineWise. (2018). Engaging Consumers in their Health Data Journey. Canberra, ACT: CHF and NPS MedicineWise. https://chf.org.au/sites/default/files/engaging_consumers_health_data_report_updated.pdf.
49. Research Australia. (2020). Public Opinion Poll on Health and Medical Research and Innovation 2020. <https://researchaustralia.org/reports/public-opinion-polling-2/>.
50. Aitken, M., de St. Jorre, S., Pagliari, C., Jepson, R., & Cunningham-Burley, S. (2016). Public responses to the sharing and linkage of health data for research purposes: a systematic review and thematic synthesis of qualitative studies. *BMC Medical Ethics*, *17*(1), 1–24.
51. Department of the Prime Minister and Cabinet. (2019). Data Sharing and Release Legislative Reforms Discussion Paper. Canberra: Commonwealth of Australia. Available from <https://www.datacommissioner.gov.au/resources/discussion-paper>
52. Department of the Prime Minister and Cabinet. (2020). Data Availability and Transparency Bill 2020 Exposure Draft Consultation Paper. Canberra: Commonwealth of Australia. <https://www.datacommissioner.gov.au/resources/exposure-draft>.
53. Ford, E., Boyd, A., Bowles, J. K., Havard, A., Aldridge, R. W., Curcin, V., et al. (2019). Our data, our society, our health: A vision for inclusive and transparent health data science in the United Kingdom and beyond. *Learning Health Systems*, *3*(3), e10191.
54. Department of Health and Social Care. (2021). Data saves lives: reshaping health and social care with data (draft). <https://www.gov.uk/government/publications/data-saves-lives-reshaping-health-and-social-care-with-data-draft>

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.