

**GETTING READY
FOR RESEARCH USING ADMINISTRATIVE DATA:**

INSIGHTS FOR COMMUNITY LEGAL CENTRES

ACKNOWLEDGEMENT OF COUNTRY

SMLS acknowledges the Bunurong People and the Wurundjeri People, the Traditional Owners of the land where we work, and First Peoples language groups and communities across Victoria and Australia. We pay our respects to Elders past and present.

We celebrate the people, traditions, culture and strength of Aboriginal and Torres Strait Islander peoples, and the fight for survival, justice and country that has taken place across Victoria and Australia.

We thank the Traditional Custodians for caring for Country for thousands of generations.

SMLS recognises the ongoing impact of colonisation, dispossession and racism experienced by Aboriginal peoples. As a Community Legal Centre, we acknowledge the violence of Australian law and its ongoing role in processes of colonisation. We recognise that sovereignty was never ceded, and that this always was and always will be Aboriginal land.

ACKNOWLEDGMENTS

The authors would like to acknowledge and thank the following who provided valuable guidance and input into this project:

- Ashleigh Newnham, Kristen Wallwork and all other staff who participated in project consultations at South-East Monash Legal Service Inc
- Dr Hugh McDonald, Victoria Law Foundation
- Nick Shearer and Ying Zheng, Fair Work Ombudsman
- Jacqui Siebel, Justice Connect
- Dr Jacqueline Weinberg, Catherine Zhou and Adam Reisner, Australian Centre for Justice Innovation, Faculty of Law, Monash University

Funding for this project was provided by a Knowledge Grant from the Victoria Law Foundation.

ACRONYMS

ACJI	Australian Centre for Justice Innovation
CLC	Community legal centre
FWO	Fair Work Ombudsman
SMLS	South-East Monash Legal Service Inc



Funded by
**Victoria Law
Foundation**

CONTENTS

Overview	4
Key Messages	5
Using Administrative Data For Research in CLCS	6
1.1 Introduction	6
1.2 What is administrative data	6
1.3 What kind of administrative data is generated by CLCs?	8
1.4 Our project – the civil justice problem of underpayment	9
The Research Process: Collaboration, Data and Ethics	10
2.1 Introduction	10
2.2 Building a research collaboration	10
2.3 Assessing the administrative data resources for the research	12
2.4 Learning through data collection	15
2.5 Navigating research ethics and consent processes	16
Reflections, Recommendations and Conclusions	18
3.1 Introduction	18
3.2 The benefits of engaging with administrative data for research in CLCs	18
3.3 Opportunities for enhancing organisational data practices	19
3.4 Improving data quality	20
3.5 Client consent and research ethics infrastructure	21
3.6 Developing a sustainable research collaboration to advance access to justice	22
Annexure 1: Administrative Data Research Checklist	23
Annexure 2: Sample List of Variables And Coding Guide Extract	28
References	31

OVERVIEW

Administrative data refers to information and records collected and stored in the everyday operation of organisations (McDonald et al, 2020). The ready availability of administrative data means that is increasingly recognised as a valuable resource for understanding the characteristics, experiences and outcomes of clients and services in civil justice contexts. In 2020, the Victorian Law Foundation (VLF) published a landmark report exploring administrative data and the way it is used in the Victorian legal assistance sector (McDonald et al, 2020). This work has made a significant contribution to our understanding of the use and potential of administrative data in the state's legal assistance sector. For Community Legal Centres (CLCs), administrative data has great potential as a means to understand and enhance services provided to clients and the community. There is also growing interest in this underutilised resource for its potential to answer questions about access to justice and the operation of legal systems.

Despite this potential, resourcing and capability barriers can limit the extent to which CLCs can make use of this data. 'Worked examples' and shared lessons from how CLCs go about research and evaluation are rare and can provide useful guidance for organisations who may engage in this kind of work in the future (see, for example, Victoria Legal Aid, 2016; Rygiel, 2021; Jones et al, 2022). This guidance can include how to go about developing research questions that can be addressed by administrative data, understanding data resources and their limitations and navigating research ethics and client consent for the use of data. Sharing these learnings can, in turn, contribute to capacity building to strengthen research and evaluation in the sector.

In 2020, South-East Monash Legal Service (SMLS) was awarded a VLF Knowledge Grant to undertake research on wage and entitlement underpayment and civil justice, using SMLS's administrative data. The research was conducted in collaboration with the Australian Centre for Justice Innovation in the Faculty of Law at Monash University. The aims of the project were to:

- Investigate the characteristics of our clients who had experienced unpaid wages and/or entitlements, and the associated employers;
- Identify the nature, range and duration of services provided to affected clients;

- Explore the outcomes legal action achieves (including the extent to which clients are recovering money through the civil justice system);
- Reveal the realities and complexities of legal processes seeking recovery of wages and/or entitlements; and
- Build and consolidate capacity in SMLS to undertake research using the service's administrative data resources.

In this briefing, we draw on our experiences over the course of the project and document what we have learned about the research process. The aim of this briefing is to share our learnings so that we may inspire and enable other community legal centres ('CLCs') and like organisations to investigate how their own administrative data could contribute to research and understanding. By providing a blueprint of our research process, we seek to contribute both to capacity building within our own organisation and others in the legal assistance sector who are interested in using their data for research.

Part 1 of this briefing provides an overview of administrative data in the CLC setting and explains how it might be used for research. It also introduces our research project, through which we generated the insights for CLCs presented in this briefing.

Part 2 explains key features of the design of the research, including developing collaborations, assessing the available data resources (explaining how we assessed, selected, gathered and analysed the data) and considerations relevant to human research ethics in this kind of research.

Part 3 reflects on our key findings from the research design and process, the potential barriers affecting efforts to do better and how these lessons might be carried forward to inform future research in CLCs. We make recommendations based on our experience for other organisations considering making more extensive use of administrative data for research.

Annexure 1 provides a brief Administrative Data Research Checklist for CLCs to consider when undertaking research using administrative data.

Annexure 2 provides examples from our work of lists of variables and the coding handbook we used in the data collection process.

KEY MESSAGES

- When planning research using administrative data, consider the range, quality, benefits and limitations of the data available to your organisation. The VLF's 2020 report on administrative data in the Victorian Legal Assistance sector is a terrific starting point.
- Case files may be a useful source of data to answer research questions of interest to your organisation in a more systematic way than the use of isolated case studies or reporting systems data (such as the Community Legal Assistance Service System). Combining multiple sources of data may be particularly fruitful.
- Administrative data can provide a cost-effective means of starting to investigate areas of service design and delivery and access to justice that have been under-explored.
- In planning your research, consider the benefits of different partners you might engage with. Partnering with a university on research using administrative data can have a range of mutual benefits for CLCs and researchers (Jones et al, 2022). We found these include capacity building, improving understanding of one another's work, developing data resources, building sustained partnerships and generating new insight into the legal assistance sector.
- The accuracy of administrative data is one of its chief challenges (McDonald et al, 2020).

Consulting with staff in your organisation who generate administrative data is an important early step in shaping your research. It can also illuminate new paths for inquiry.

- The scope of your research may need to be shaped by feasibility constraints associated with the way data is held and understood in your organization (including over what period and in what format).
- Your organisation has deep expertise about your own data and the types of legal problems you are dealing with. This expertise has an important role to play in research using administrative data. Involving your staff in data collection is a valuable way to make the most of this expertise in research and build capacity for future projects.
- Using a pilot in your data collection is a useful way to check the feasibility and accuracy of your research approach and implement any changes needed before completing the entire data collection.
- At the outset and throughout your research, you should consider ethics and consent in relation to your use of administrative data. Depending on your project partner/s, you may be able to access established research ethics infrastructure for your research (such as that found in universities).

PART ONE

USING ADMINISTRATIVE DATA FOR RESEARCH IN CLCS

1.1 INTRODUCTION

This part briefly introduces the nature of administrative data, the range of data held by CLCs like SMLS and the way it might be used for research and evaluation. It draws extensively on the VLF's detailed investigation of these issues in Apples, Oranges and Lemons: The use and utility of

administrative data in the Victorian legal assistance sector (McDonald et al, 2020), an invaluable starting point for organisations seeking to improve their work in this domain. In this part we also describe our research project (the case study used to explore administrative data in this briefing).

1.2 WHAT IS ADMINISTRATIVE DATA

The benefits of administrative data in research contexts mostly flow from the fact that it is existing data.

Administrative data is information collected and stored as part of the everyday functions of organisations, rather than for a specific research purpose (McDonald et al, 2020; Broadhurst et al, 2021; McLennan, 2018). There is growing interest in this underutilised resource for its potential to answer questions about access to justice and the operation of legal systems, and create an evidence base to inform policy and advocacy work. We also see strong potential in this data to inform questions of interest to CLCs about service design and delivery.

The benefits of administrative data in research contexts mostly flow from the fact that it is existing data. As a result, analyses using this data is often low cost, particularly compared with primary data collection (such as using surveys, interviews or focus groups) (McLennan, 2018). Research using administrative data may also be

faster than other research approaches. Service user or participant burden is commonly minimal or non-existent when administrative data is used (Statistics Canada, 2017), and may enable the experiences of clients or groups who are otherwise unlikely to participate in research to be explored (McDonald et al, 2020). Additionally, administrative datasets are commonly longitudinal, in that they have been collected over time, and may therefore allow changes in service provision and outcomes to be monitored (McDonald et al, 2020).

Nonetheless, there are important challenges and limitations associated with using administrative data for research and evaluation. Firstly, because the data has been collected for non-research purposes, it may not address research questions of interest. The data is confined to the scope of the information that is routinely

collected, and it may not be possible to supplement the data in useful ways, unless some form of linkage is feasible. Most significantly, a range of data quality issues may compromise effective use of administrative data.

These can include:

- Inaccurate and incomplete data collection;
- Differing local priorities, which may make some parts of the data more or less accurate or complete; and
- Inconsistency in the meaning of data collected over time or by different data contributors (eg different branches of one organisation, or individuals responsible for data entry) (McDonald et al, 2020; Statistics Canada, 2017; McLennan, 2018).

There may also be costs and delays associated with reviewing and cleaning administrative data to make it suitable for analysis. By its very nature, administrative data may fail to capture the complexity, depth and value of services delivered (McDonald et al, 2020). This can prove particularly problematic in the CLC setting, where

services are often provided in a context of considerable client disadvantage and multiple legal and other problems (Coumarelos et al, 2012). These benefits and challenges associated with administrative data are all important considerations for CLCs embarking on research and evaluation using these data sources.

SMLS has significant administrative data resources by virtue of our long history in the community legal sector. In the past, however, we have typically used case studies of individual client experiences and matters to support our advocacy, rather than looking across cases or clients in our broader data. The chief barrier we have encountered is a common one in CLCs – namely the resources to undertake the required capacity building work, which necessarily competes with our work providing frontline legal assistance to our clients and community (McDonald et al, 2020).

When planning research using administrative data, consider the range, quality, benefits and limitations of the data available to your organisation. The VLF's 2020 report on administrative data in the Victorian Legal Assistance sector is a terrific starting point.

1.3 WHAT KIND OF ADMINISTRATIVE DATA IS GENERATED BY CLCS?

In the day-to-day functioning of SMLS's service, we generate and accumulate a large volume of data regarding our clients and the services we provide.

CLCs routinely collect demographic information about their clients, the nature of clients' legal problems and the services clients are provided with (McDonald et al, 2020). As the VLF has established, CLCs in Victoria make use of a range of data systems to gather and report data about clients, types of legal problems and services delivered.

In the day-to-day functioning of SMLS's service, we generate and accumulate a large volume of data regarding our clients and the services we provide. Our organisation principally uses the Community Legal Assistance Service System (CLASS), which is used to facilitate reporting about the performance of our service (McDonald et al, 2020). McDonald et al categorise the administrative data collected using this and other like systems as falling into three main categories, namely data on:

- Service users (including social and demographic characteristics, such as age, gender, income, education and employment status);
- Service and legal matter type; and
- Referral and secondary consultation.

In addition to the administrative data generated for service reporting, a considerable wealth of data is produced by CLCs in the course of work on files with individual clients. Case files are themselves an untapped source of administrative data, in the form of such materials as file notes, correspondence (with clients, opposing parties and others) and court documents. Case files contain a far richer range of information about the profile of our clients, the legal problems they have experienced and the nature of outcomes of the services we have provided than it is feasible to obtain through CLASS. Importantly, however, while it is possible to rapidly generate reports in CLASS and other like systems to explore the high-level characteristics of samples of clients or matter types, working with case files requires a more labour-intensive and time-consuming approach. In our research on the civil justice problem of unpaid wages and/or entitlements, we combined these two sources of data (CLASS and case files) to develop a profile of the clients, services and outcomes involved.

Case files may be a useful source of data to answer research questions of interest to your organisation in a more systematic way than the use of isolated case studies or reporting systems data (such as CLASS). Combining data may be particularly fruitful.

1.4 OUR PROJECT – THE CIVIL JUSTICE PROBLEM OF UNDERPAYMENT

Our project generated a profile of the civil justice problem of employment underpayment through analysis of administrative data about clients, services and outcomes in these actions at SMLS.

In recent years there has been considerable media and law reform attention to the growing problem of wage and entitlement underpayment. The focus has mostly been on civil and criminal penalties against wrongdoing employers, with little attention to the outcomes of that flow to underpaid workers and contractors through the civil justice system.

Our project generated a profile of the civil justice problem of employment underpayment

through analysis of administrative data about clients, services and outcomes in these actions at SMLS. We undertook the project with the aim of informing better client service, advocacy and data practices whilst building sustainable research capacity at SMLS.

Administrative data can provide a cost-effective means of starting to investigate areas of service design and delivery and access to justice that have been under-explored.

PART TWO

THE RESEARCH PROCESS: COLLABORATION, DATA AND ETHICS

2.1 INTRODUCTION

This part outlines the work we undertook to:

- Foster a collaborative relationship between partners in the research;
- Assess the administrative data available for the project; and
- Navigate the human research ethics review system.

2.2 BUILDING A RESEARCH COLLABORATION

ACJI's research provides evidence for effective improvements to law and justice systems through empirical research. A core part of ACJI's work is building capacity for empirical civil justice research through training and mentoring of students, staff and research partners.

SMLS is a community legal service based in the southeast region of Melbourne. We have a long history of providing free legal advice and assistance to community members experiencing vulnerability and disadvantage. We practice in a broad range of areas of law, including employment law. In 2019 and 2020, SMLS provided legal assistance to over 12,700 clients. Of those clients, more than 1600 (around 13 per cent) sought assistance in relation to an employment law matter.

As explained above, SMLS has accumulated a significant asset in the form of administrative data, but to date has not optimised its data to inform service design and delivery and policy advocacy work. Historically, the chief barrier SMLS has encountered is funding and investment in people and time to undertake the required capacity-building work. This is a common problem in the community legal sector, where resources are stretched. McDonald et al (2020) highlighted that partnerships between legal services and universities (like that developed in this

project) can be a way to improve and enhance the use of administrative data. Using a research collaboration to strengthen data capacity at SMLS will position the organisation to improve its service delivery to clients and advocacy work, as well as sharing its insights and improved practices with its partners and other Community Legal Centres.

In this project, SMLS partnered with the Australian Centre for Justice Innovation ('ACJI') at Monash University to explore the extent to which our administrative data could be used for research purposes. ACJI is a research centre in the Monash Faculty of Law. ACJI conducts high quality research and evaluation to support effective civil justice policy and practice innovation. ACJI's research provides evidence for effective improvements to law and justice systems through empirical research. A core part of ACJI's work is building capacity for empirical civil justice research through training and mentoring of students, staff and research partners.

Over the course of the project, project investigators at ACJI and SMLS attended regular meetings approximately once per month. A key focus of these meetings was for ACJI to support and mentor SMLS through the stages of the project. Initial meetings focused on discussion on the purpose and design of the project and the respective roles of the partners, with later meetings concentrating on developing a shared understanding of the data and exploring the interpretation of the findings. These lively meetings provided valuable opportunities to 'brainstorm' ideas and discuss broader access to justice issues which may have relevance to further research or policy work. They were a critical part of establishing the rapport between key investigators, further strengthening the partnership. This was achieved notwithstanding the partners predominantly attending these meetings remotely.

To measure the success of the capacity-building aims of the project, we conducted a

post-project consultation with key investigators from both SMLS and ACJI. We identified that the collaboration generated mutual benefits for the partners including through access to training, development and mentoring opportunities; increased understanding of the value of collecting data and appreciation for attention to data quality and accuracy; novel research data and findings; and the development of a joint approach likely to be carried forward in future research projects and outputs.

In planning your research, consider the benefits of different partners you might engage with. Partnering with a university on research using administrative data can have a range of mutual benefits for CLCs and researchers (Jones et al, 2022). We found these included capacity building, improving understanding of one another's work, developing data resources, building sustained partnerships and generating new insight into the legal assistance sector.

2.3 ASSESSING THE ADMINISTRATIVE DATA

RESOURCES FOR THE RESEARCH

We engaged in a multi-stage process of identifying relevant cases, collecting data about those cases (from CLASS and case files), merging the two sets of data and cleaning and analysing the merged dataset to generate findings.

One of the key areas of work for the project was building our understanding of the administrative data resources available to us and how they might be used to address our research questions about wage and entitlement underpayment matters dealt with by SMLS. We engaged in a multi-stage process of identifying relevant cases, collecting data about those cases (from CLASS and case files), merging the two sets of data and cleaning and analysing the merged dataset to generate findings. This process is summarised in Figure 1 and explained further below.

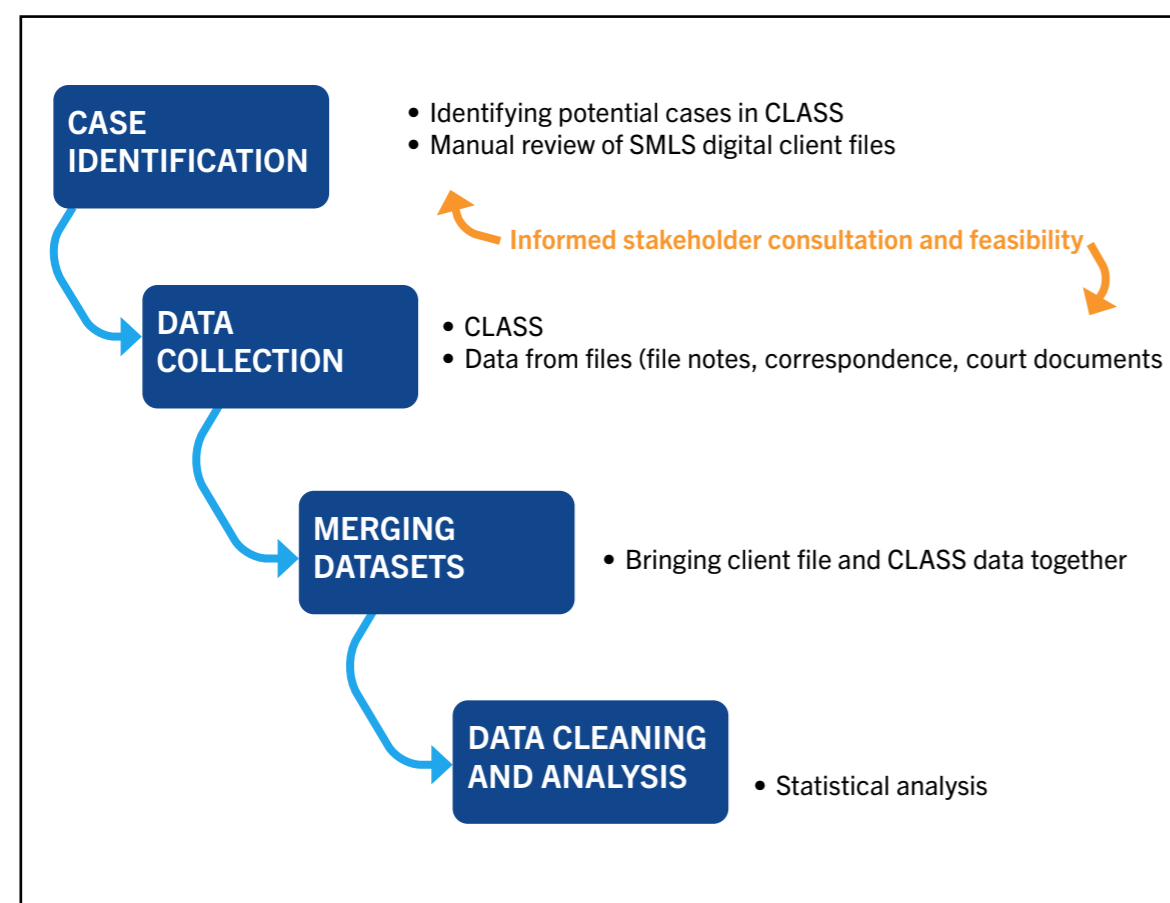


Figure 1: Overview of data identification, collection and analysis

The first steps were identifying those client matters that were relevant to the research project and the nature of the available data. To assist with these tasks, we consulted with SMLS staff about their data-related practices and understanding. Early in the project, we conducted a group consultation session with 14 SMLS lawyers and followed this up with one-on-one interviews with four lawyers who practised in employment law. The one-on-one interviews took up to 30 minutes.

Through these consultations, we sought staff input and feedback on:

- Areas of interest regarding wage and entitlement underpayment matters the project could pursue;
- Lawyers' insights into the characteristics of underpayment clients and non-payers;
- The possible client and legal pathways for debt recovery for unpaid work (including lawyers' observations about client pursuit of payment through VCAT or court, and barriers to recovery);
- Lawyer data practices and experiences, including use of CLASS for relevant clients (for example, the matter type/s used for employment underpayment claims) and recording of client outcomes; and
- Lawyer referral practices.

The results of these consultations enabled us to establish which data we should extract from CLASS and the most relevant parts of the client files for obtaining information relevant to our research. Some of these considerations related to issues of data quality, in that we needed to be sure that the data categories we used in CLASS were understood and applied accurately at the stage of data input (see McDonald et al, 2020). For example, we had initially intended to explore both underpayment experienced by employees

and debt recovery proceedings involving independent contractors; we found, however, that we identified fewer matters involving independent contracts than what we expected in CLASS, and SMLS staff used a range of different problem type categories in relation to these matters, describing them either as a 'credit or debt owed to client' matter or an 'employment conditions entitlements' matter. The category 'credit or debt owed to client' was a broad category encompassing all debt matters, rather than being specific to unpaid independent contractors. Accordingly, it was not feasible to focus on independent contractors in our research because of the difficulty of locating those matters accurately. A summary of the key questions we used in interrogating this issue in the form of a. Administrative Data Research Checklist is provided in Annexure 1.

The accuracy of administrative data is one of its chief challenges (McDonald et al, 2020). Consulting with colleagues who generate administrative data is an important early step in shaping your research. It can also illuminate new paths for inquiry. More broadly, attention to quality, accurate, and consistent data practices is likely to generate significant benefits for future research and analysis in your organisation.

Feasibility, time and resourcing constraints also played a role in shaping our data collection, as we did not have capacity to analyse all of the SMLS files over an extended period. In view of the COVID-related access challenges during the period in which our research was conducted (with social distancing and lockdowns limiting our access to files), it was necessary for us to limit our research scope to client files that were available digitally. SMLS had moved to a paperless system in early 2020 so the majority of our client file data was readily available in digital form.

The scope of your research may need to be shaped by feasibility constraints associated with the way data is held and understood in your organization (including over what period and in what format).

Ultimately, we determined that we would focus our research on CLASS data and client files for clients who were identified as having first contacted SMLS regarding their work-related debt recovery matter between 1 January 2019 and 31 December 2020. We included client matters where we assisted on an ongoing basis as well as matters where we provided discrete legal assistance, such as one-off advice. These clients had matters with the problem type 'Employment conditions entitlements' in CLASS.

Based on staff consultation, we also identified that clients were from time to time referred to the Fair Work Ombudsman ('FWO') and Justice Connect for further assistance. In some instances, clients were directed to courts and tribunals to represent themselves in proceedings. Accordingly, we sought supplementary data from VCAT, Justice Connect and the FWO, in particular whether the client was able to recover their debts for unpaid work. We also sought feedback from the FWO and Justice Connect on the parameters of our research project.

2.4 LEARNING THROUGH DATA COLLECTION

SMLS investigators underwent training within their organisation on extracting demographic and service data for relevant clients and matters.

An important part of the capacity-building design and intent of our project was that the data collection was undertaken by the SMLS investigators. The expertise and understanding of these investigators in the subject matter of the project was crucial for ensuring the most relevant data was used in the research: CLCs are experts on the strengths, weaknesses and quality of their own data, and have invaluable insight on data-related practices (McDonald et al, 2020; see also Buhler & Kaiser-Derrick, 2020). Some study data could be readily obtained from CLASS. SMLS investigators underwent training within their organisation on extracting demographic and service data for relevant clients and matters. They drew on this training in compiling a range of demographic and service data for the client population of interest for the research.

Your organisation has deep expertise about your own data and the types of legal problems you are dealing with. This expertise has an important role to play in research using administrative data. Involving your staff in data collection is a valuable way to make the most of this expertise in research and build capacity for future projects.

To collect the data held within individual client files, the research team engaged in an iterative process of developing a coding handbook and spreadsheet. The coding handbook contained agreed definitions of key variables, drawing on the staff consultations and project team expertise. The handbook was relied on by the SMLS investigators when reviewing the client files and

extracting relevant data. The spreadsheet used drop-down menus of pre-defined responses where possible rather than permitting free text entry, to minimise errors. A sample list of variables together with illustrative extracts of the coding guide are provided in Annexure 2.

Two SMLS investigators (members of the research team) conducted the review of the client files and associated data collection. The SMLS investigators reviewed each client file identified as being possibly relevant to the study on the basis of the CLASS data and limited the review to those files that were available digitally. As part of their training and the development of the coding guide and data collection spreadsheet, the SMLS investigators conducted a pilot on a sample of client files to test whether they independently extracted the same data from identical files. Findings from the pilot were presented and discussed in a project meeting. Learnings from this process were implemented as refinements to the coding guide and spreadsheet, which were then used in the final data collection. Once data collection was complete, ACJI then merged the CLASS and client file review datasets, undertook data cleaning and performed statistical analysis using STATA software (StataCorp, 2019).

Using a pilot in your data collection is a useful way to check the feasibility and accuracy of your research approach and implement any changes needed before completing the entire data collection.

2.5 NAVIGATING RESEARCH ETHICS AND CONSENT PROCESSES

Engaging in ethical research and evaluation may involve establishing bespoke processes for client consent outside of formal institutional ethics committees (see eg Szczepanska & Blomkamp 2020).

Being attentive to questions of client consent for the use of data is critical to conducting research with administrative data in an appropriately ethical way. Staff in the Victorian legal assistance sector identify engaging with ethics approval processes as one of the difficulties associated with their use of administrative data (Jones et al, 2022) and collaboration with academics on research (McDonald et al, 2020). Engaging in ethical research and evaluation may involve establishing bespoke processes for client consent outside of formal institutional ethics committees (see eg Szczepanska & Blomkamp, 2020). Where CLCs partner with organisations (such as universities) with established human research ethics infrastructure, they may instead engage with those processes and draw on the experience of their partners. There are benefits to obtaining such institutional ethics approval prior to engaging in research. Ethical approval processes require research designs, data, consent and associated risks to be closely considered and articulated, and prompt research teams to be mindful of treating client data with optimum care and respect. Ethics approval also opens opportunities for the research to be published in scholarly outlets, contributing to the body of evidence on a particular subject matter. Engaging with ethics approval processes may also shed light on how an organisation's practices around consent and data might be improved to facilitate future research.

At the outset and throughout your research, you should consider ethics and consent in relation to your use of administrative data. Depending on

your project partner/s, you may be able to access established research ethics infrastructure (such as that found in universities).

At the outset of our research it was clear that ethics approval would be required for both the CLASS data and client file analysis. At the time of undertaking this research, SMLS did not have a routine process of obtaining clients' consent to use their data for research purposes. The collaboration with ACJI gave SMLS access to the human research ethics governance framework at Monash University as well as support navigating the approval process. The application required us to address benefits arising from the research to participants and the community, and analysis of risks to participants associated with our use of the data.

We applied for and obtained a waiver of client consent for our use of the data in the research. We took this approach on the grounds of the following factors:

- We highlighted the evidence base on the use and outcomes of legal services is incredibly limited, particularly in relation to recovery of employment underpayment claims. The benefits of the project therefore outweighed any risks of harm from not seeking consent. These benefits included the potential for the research to contribute to client advocacy and access to justice issues in relation to the significant problem of wage and entitlement underpayment; the scope for the reach to contribute to

improvements in service delivery within the civil justice system; and strengthening SMLS's data-related capacity to engage in similar research in future.

- We identified that there was a very small risk that the clients of SMLS whose data is drawn on for the research would recognise their individual data in the reports produced in relation to the research and experience any adverse reaction, or that data breaches might occur. We identified a range of steps to manage this risk, including the de-identification of data, the data collection being undertaken exclusively by SMLS staff and ensuring any data shared electronic between the research team members were secure and password protected.
- There was no known or likely reason that clients would not have consented if they had been asked. In the past SMLS has on a small scale approached individual clients to draw on their experiences in developing case studies for use in advocacy campaigns and clients have been consistently supportive of the use of their information in this way.
- It was impractical for us to individually contact each of the hundreds of clients in our sample, given the volume of clients and the likelihood that the contact details for some clients would have changed since our last engagement with them.

PART THREE

REFLECTIONS, RECOMMENDATIONS AND CONCLUSIONS

3.1 INTRODUCTION

This part outlines our reflections and findings about:

- The extent to which our administrative data could be used for research purposes;
- Practice developments required to engage in further research of this kind in the future, particularly in relation to research ethics and client consent; and
- Whether ACJI and SMLS were able to develop a sustainable research partnership.

3.2 THE BENEFITS OF ENGAGING WITH

ADMINISTRATIVE DATA FOR RESEARCH IN CLCS

It can provide an evidence base to indicate the scale of an issue and reveal patterns of systemic issues related to access to justice.

We were pleased to discover that there was indeed scope to use our organisation's administrative data for research purposes. Our experience therefore lends further weight to the increasing calls for this to happen. The quantitative data can add further understanding to the client experience and complement the qualitative data that can be derived from client case studies. It can provide an evidence base to indicate the scale of an issue and reveal patterns of systemic issues related to access to justice.

Through our staff consultations and in executing the research, we identified that a range of data was collected and recorded sufficiently consistently by SMLS lawyers for the purposes of our project. For example, the key components of client files for the matters of interest were found across the files from which we collected data (such as file notes of initial client interviews, letters to client notifying them that we would be

providing ongoing assistance, a letter of demand and final letters to clients notifying them of the closure of their file). The relevant documents held within client files were identified with relative ease as there appeared to be some consistency with the way files were labelled.

We also identified a range of unintended benefits arising from the process of engaging in this research project. SMLS staff reported that the process of collecting data for the case file analysis served as a useful learning opportunity to build knowledge of the substantive law and enhance professional skills. It was also reported that it offered a valuable opportunity to encounter some good precedents that could be used in their practice.

From a longer-term service delivery perspective, the case file analysis also generated benefits in enabling us to:

- Identify training needs for staff;
- Recognise opportunities for developing consistencies amongst the lawyers in managing cases; and
- Highlight the scope for developing shared data practices with external organisations to enable more complete mapping of a client journey

From a policy and advocacy perspective, the case file analysis was also useful in more clearly identifying how these matters may interact with other areas of law, for example where a client presents with other legal issues.

3.3 OPPORTUNITIES FOR ENHANCING ORGANISATIONAL DATA PRACTICES

Through this research project, we identified a number of opportunities to enhance our Centre's data practices in order to better identify those cases which may be relevant to a research project and also to more comprehensively collect relevant data.

Getting administrative data 'research ready' requires capacity, consultation, planning and infrastructure. For administrative data to be useful for research, there needs to be consistency in what is collected, and where it is recorded, together with shared understanding of the definition of key terms and concepts.

Our staff consultation was critical to the process of identifying the cases relevant to the study and in the process of data collection. Speaking to the staff directly involved in collecting and recording the data and finding out more about their deep experience working on files and their data handling practices was crucial for locating

the data we needed for the project. Much of this insight is held the minds of individual lawyers, so the findings arising from staff consultation was invaluable. We also found it particularly beneficial to speak with staff on a one-on-one basis so that staff had the opportunity to speak candidly about their data practices.

Through this research project, we identified a number of opportunities to enhance our Centre's data practices in order to better identify those cases which may be relevant to a research project and also to more comprehensively collect relevant data.

3.4 IMPROVING DATA QUALITY

One way to reduce the possible burdens of data collection could be expanding the use of templates for critical documents that include prompts to collect key data of interest.

Consistent with observations made by McDonald et al (2020) about the wider Victorian legal assistance sector, we found in our project that some data that appears promising for research is not collected consistently across all matters. For example, in consultations staff reported that it seemed common that clients on temporary visas would opt not to pursue debts for unpaid work due to concerns about negative impacts on their visas. We sought to collect data on whether clients expressed concerns regarding their visa during our file review, but were unable to identify many instances of this in the files. It is unclear whether lawyers were consistently recording this information, or whether the issue of concerns regarding visas was always put to clients. Data of this kind needs to be consistently obtained in order to be used in research and changes to the data we record on intake might be one way of achieving this. Consultation with other organisations in the sector may also contribute to expanded data consistency on these issues.

At the same time, changes to data collection practices must be assessed in terms of the

added burdens this may create for staff and clients. SMLS staff are currently required to collect a range of personal and sensitive data from clients. Staff are often faced with dealing with a high volume of clients with significant time constraints. One way to reduce the possible burdens of data collection could be expanding the use of templates for critical documents that include prompts to collect key data of interest. In our context, for example, staff identified the following documents as places where key data was typically recorded:

- File note of the initial client interview;
- Initial letter to the client confirming the provision of ongoing assistance and the scope of that assistance;
- Letter of demand; and
- Final letter to the client confirming completion of our assistance in their matter.

These materials could be assessed to determine whether templates might be used to facilitate the inclusion of standard items to facilitate future research on topics of interest.

3.5 CLIENT CONSENT AND RESEARCH

ETHICS INFRASTRUCTURE

Engaging in ethical research and evaluation may involve establishing bespoke processes for client consent outside of formal institutional ethics committees (see eg Szczepanska & Blomkamp 2020).

Our experience demonstrates that it may be possible to receive ethics approval to use administrative data for research purposes, in circumstances where prior client consent may not have been initially obtained. It is unclear, however, whether there would be a strong basis to again seek waiver of client consent in future ethics applications and research.

We recommend further work in developing a sustainable framework for client consent to the use of administrative data in research and evaluation in our organisation and others in the sector. It would be appropriate to consult more widely with clients and other stakeholders within the sector as part of this work. In developing this framework, we anticipate there will be a need to balance a range of priorities, including:

- Ensuring that client consent is informed and freely given, particularly given concerns about whether clients will perceive that their consent is the 'cost' of accessing legal services;
- Maintaining the integrity of the data we collect; and
- Minimising the burden on the client and the legal service in relation to data collection, especially given the need to build and maintain rapport with our clients and their trust in our service and staff, especially in relation to highly sensitive legal problems.

We anticipate it may be almost impossible to comprehensively identify the nature of all future research that administrative data may be

used for. A widely-phrased blanket consent may assist with broadening the range of potential research; however, this approach may give rise to concerns about how informed client consent might be. Identifying the appropriate time to seek client consent is also challenging. If consent is obtained at the commencement of service to a client (such as through inclusion in standard intake forms), then this may contribute to the completeness of the data. There is a risk, however, that clients may feel pressured to give that early consent if they feel withholding it may affect access to legal help or the quality of the assistance they could receive. If consent is sought at the conclusion of the service, this may compromise the completeness of the data if we lose contact with a significant proportion of clients prior to the conclusion of their matter.

We recommend further consideration be given as to measures that can be put in place to ensure the client does not feel coerced into providing the consent and that the client is assured that the granting or withholding of that consent will have no bearing on the access to service or quality and content of the assistance provided. One possibility is considering whether these risks may be mitigated by seeking that consent in the absence of the lawyer who will be directly assisting the client. For example, it may be that another staff member within the legal service seeks the client consent, and the lawyer providing assistance is not made aware of whether that client has consented. Consideration would need to be given to the administrative and resourcing implications of such an approach.

3.6 DEVELOPING A SUSTAINABLE RESEARCH COLLABORATION TO ADVANCE ACCESS TO JUSTICE

It was also helpful to have an early meeting from the outset to clarify the terms of the research project, the roles and responsibilities a proposed timeline with key milestones.

We conducted an informal post-project consultation to explore the learnings for our organisations from our project, documenting key themes in the respective experiences of SMLS and ACJI.

SMLS staff reported:

- Enhanced appreciation of administrative data and increased confidence and capacity to engage with it for research purposes;
- Increased confidence and capacity to approach research institutes such as the ACJI with research project proposals;
- Improved understanding of the level of effort and investment required to engage with administrative data for the purposes of research; and
- Greater understanding of research design and methodology and its importance to research projects.

ACJI staff reported:

- Increased understanding of how the Victorian legal assistance sector and our organisation operate;

- Enhanced appreciation of the challenges and burdens that a CLC may face in data collection and analysis;
- That seeing how research could contribute to advocacy was immensely rewarding; and
- Gratitude associated with the opportunity to access novel data and understanding about the operation of Victoria's civil justice system, both through the administrative data and practice expertise of SMLS and its staff.

The support and mentoring offered by the ACJI was invaluable to the success of this project. Through the regular training provided by ACJI, SMLS staff reported increased confidence in engaging with administrative data for research purposes. The strength of the partnership was helped also by having regular well-structured meetings with clear points arising from each meeting. It was also helpful to have an early meeting from the outset to clarify the terms of the research project, the roles and responsibilities a proposed timeline with key milestones.

PART FOUR

ANNEXURE 1: ADMINISTRATIVE DATA RESEARCH CHECKLIST

4.1 WHO IS THIS CHECKLIST FOR?

This checklist is aimed at Community Legal Centres and other organisations who wish to engage in research using their administrative data. It draws on our experience of doing this,

as well as published data quality frameworks and assessment tools (ABS, 2009; AIHW, 2019; WGEA, 2017; Smith et al, 2018; Iwig et al, 2013).

4.2 WHY IS THIS CHECKLIST USEFUL?

Administrative data is information collected and stored as part of the everyday functions of organisations. There is a wealth of possibility in using this data in CLCs to better understand clients, services and outcomes. It may, for example, offer a low-cost evidence base to support advocacy on access to justice issues and inform improvements to service delivery. Care must be taken, however, to ensure that the data is of sufficient quality and that its meaning and limitations are understood.

The checklist provides a series of questions or prompts for you to consider throughout your research project about administrative data you might use. It is intended to help you understand

the data you are working with and to improve the quality of your research. It is not an exhaustive list of considerations, but a place to start. Although it is presented as a sequence, it need not be used only in that way – you might find it useful to revisit earlier stages of the checklist as your work progresses.

Importantly, this checklist does not provide a guide to undertaking analysis of the data you collect. The checklist is best used by staff involved in developing research questions and identifying, assessing and collecting administrative data for research. Ideally, this will involve close collaboration with the team members undertaking the analysis of the data.

4.3 ADMINISTRATIVE DATA RESEARCH CHECKLIST

STARTING POINTS

1. What are your research questions?

Define your research questions. This can start off broad. The research questions could be confined to a particular area of law or issue; focus on certain cohorts of clients; or cover particular time periods and/or geographical locations. For example, you might want to explore the demographic profile of clients presenting with a particular type of legal problem over a specific time frame.

2. Does the available data match your research questions?

You must ensure that the administrative data you intend to use properly matches your research question/s – that is, will your data answer the research question/s you have defined? You may need to review this match multiple times throughout the life of your project, as your sense of the quality of the data, the focus of your research and its feasibility become clearer.

3. Have you considered ethics and client consent for your use of the data?

Using client data for research requires engaging with questions about whether clients have consented to the use of their data, or would be likely to. You should consider these issues before you begin accessing, collecting and analysing data.

Review your existing organisational policies and procedures regarding client consent for use of data. Is the research you are undertaking consistent with these policies and the consent clients have provided? Are these policies sufficient for the work you are undertaking? Are there models of ethics governance you might borrow from? You should consider how you will be using the data, how you intend to publish your research findings and whether there is any risk to clients posed by the research process or publication of the findings.

You could consider partnering with a researcher or research organisation to conduct the research project. They may assist with capacity building in your organisation, and may enable you to access the organisation's research ethics governance processes.

IDENTIFYING AND ASSESSING DATA

4. What sources of data do you plan to use?

Identify relevant sources of data within the organisation. This may include client files and data held within a client data management system. In working out what data to collect, think about:

- The relevance of the data to the research questions;
- The likely quality of the data (for example, whether it is collected routinely, accurately and completely); and
- Feasibility – that is, the effort and time needed to locate and collect the data in view of its importance to the research questions and the resources available to you for the project (eg time, personnel, funding).

You might also consider whether it is possible to link multiple sources of data (eg data from client files and data from a client data management system).

5. Consider data limitations and quality

Administrative data can be cost-effective and accessible, but it also has important limitations. It is limited to the data that is routinely collected in your organisation, and the collection is historical – this means in some cases it may not be possible to supplement it with data that better fit for your research questions. In other cases – and where resources allow – collecting additional 'snapshot' data may be useful to enable you to extrapolate on the administrative data that is available to you.

The utility of your organisation's administrative data may also be limited by its quality. The Australian Bureau of Statistics identifies seven dimensions of the quality of data: institutional environment, relevance, timeliness, accuracy, coherence, interpretability and accessibility (ABS, 2009). You should consider whether the data you plan to use are of sufficient quality for your research to be sensible, feasible and reliable. The following kinds of questions might assist in making that assessment:

- Who collected the data and for what original purpose?
- At what stage in the duration of a client's involvement with your organisation is the data collected and/or reviewed?
- Is the data consistent (for example, does the data require a lot of judgment from the person responsible for inputting it, and have standard classifications been used in the data collection)?
- Over what period does the data extend (ie what are the earliest and latest dates covered by the data)?
- How complete is the data? How much data is missing?
- Is the data available electronically or in hard copy?
- Have there been any major organisational or system changes during the time the data has been collected that might have affected its quality? These might include changes in in service funding and capacity; service models; eligibility rules; and significant external events that may have affect service provision (such as the COVID-19 pandemic).

6. What can your organisation's staff tell you about the data?

Consulting with the staff who are involved in collecting and recording the data relevant to the research is an important way to generate insight into the quality of the data and the feasibility of your planned research. These consultations could be used to explore organisational data collection practices and staff feedback on the data you intend to draw on (for example, whether there are additional or alternative sources of data, whether data is collected reliably, the shared understanding of key variables, the range of data staff are involved in collecting and where and how they gather and record data).

7. Define the variables of interest for your research

It is important to ensure there is shared understanding in your research team of the meaning of variables, particularly where multiple team members are involved in collecting data. Some variables, like dates of birth, are uncontroversial and are clearly understood. In other cases, however, you may need to take steps to ensure that you have defined and documented the meaning of a variable so data is collected consistently and that definitions are well understood within the organisation and the research team.

If using a client data management system, is there a data dictionary that explains what data means? If so, use it. If reviewing client files to collect data, you should develop a guide to document the meaning of the variables you are collecting data on, the possible values you will record (eg 'yes', 'no', a number or a date) and where the information should be located. It is good practice to undertake a pilot of your data collection using the guide, which will enable you to test its effectiveness (ie whether data can be collected consistently), and make improvements to the guide where necessary.

8. Establish how data will be collected

If collecting data from client files, you should set up a consistent way to record the data. This could involve a spreadsheet, or an online interface. Be guided by the approach that best fits the analytical approach you will be using (for example, any statistical analyses that might be planned). If using a spreadsheet, it may be prudent to limit the possible responses to a drop-down menu rather than using free text for the sake of efficiency and to limit typographical errors.

INTERPRETING FINDINGS AND REFLECTING ON YOUR EXPERIENCE

9. Interpreting findings

When interpreting the findings, consider the nature of the data you have used in your research and whether the findings seem consistent with your observations and experience in your organisation. This is not about 'second guessing' the research, but rather checking and supplementing the findings with your practice-based expertise. Your organisation's expertise has a valuable role to play in understanding the findings of research using administrative data.

10. What might you do differently next time?

Consider whether your experience working with administrative data prompts any recommendations to enhance your organisation's data

practices. Was any data consistently missing? Would this data be valuable for research purposes? Are there any hurdles for staff to collect and record this data? Are there any measures or training that could be implemented to mitigate any burdens associated with collecting and recording this data, such as developing templates? By engaging in this kind of reflective practice, you can pave the way for future projects using administrative data in your organisation.

PART FIVE

ANNEXURE 2: SAMPLE LIST OF VARIABLES AND CODING GUIDE EXTRACT

A. SAMPLE LIST OF VARIABLES

i. Client characteristics

- Date of birth
- Gender
- Country of birth
- Aboriginal and Torres Strait Islander status
- English language proficiency (written, spoken)
- Interpreter required
- Main language spoken at home
- Year of arrival in Australia
- Number of dependent children
- Number of other dependents
- Employment status
- Centrelink status
- Income
- Education level completed
- Job title
- Job industry

ii. Claim characteristics

- Presence of other legal issues and issue type
- Start date of employment

- End date of employment
 - Gross total amount of client's claim
 - Record on file of an amount being recovered
 - Gross total amount recovered by the client
- iii. Nature, range and duration of services
- Date of file opening
 - Date of file closure
 - Presence of letter of demand on file
 - SMLS attended an ADR event with the client
 - SMLS drafted court documents
 - SMLS on record at court as acting on behalf of the client
 - If legal proceedings commenced, date initiating application was filed
 - If legal proceedings commenced, date of final orders
 - If legal proceedings commenced, forum of proceeding
 - SMLS assisted with enforcement proceedings
 - Closed file as lost contact with client

B. CODING GUIDE EXTRACT

'We developed a coding framework to guide and support systematic and consistent data extraction and collation from the case files. Below we present an extract of this coding guide.'

Column	Variable and definition	Response options	Documents reviewed in the client file
<i>Variables relevant to client and non-payer profile</i>			
G	Client's job title If not stated, leave blank	Free text	Letter of demand Request for employee records Opening letter Close letter Earliest dated file note conversation with client
J	Did client have other employment-law related issues with the employer? Include any queries not related to entitlements. May include a dismissal related claim, complaints of bullying, complaints of discrimination, workplace injury, unpaid superannuation and/or a general protections claim.	Yes No	Letter of demand Opening letter Earliest dated file note of conversation with client
L	Is there a reference to bullying in the workplace?	Yes No	Opening letter Earliest dated file note conversation with the client
V	Gross total amount of client's claim If only an estimate of the claim is given, type the estimated amount. If client has only partially quantified their claim, type here amount that has been partially quantified. For example, if client has quantified unpaid wages but not the unpaid super owing, type in amount client has given in unpaid wages. If not clear whether amount claimed is gross or net, type amount as appears in the file. If the client has not provided any indication of the amount of their claim, leave blank.	Free text, type amount	Letter of demand

Variables relevant to nature, range and duration of services			
W	Date of opening letter Leave blank if no opening letter on file	Free text, type using format DD/MM/YYYY	Opening letter
X	Date of close letter Leave blank if no closing letter on file	Free text, type using format DD/MM/YYYY	Close letter
Y	Is there a letter of demand on file?	Yes No	Letter of demand
AG	Record on file of there being a recovery Select yes if after receiving advice or assistance from SMLS, there is a record on file of the client confirming that payment was received from the other party. Include both full and partial payment of what the client had expected to receive. Otherwise, select no.	Yes No	Close letter Earliest dated file note conversation with client
AH	Gross total amount recovered by client Record amount client instructed had actually received. Leave blank if: <ul style="list-style-type: none"> The client did not recover any amount Amount recovered is not recorded in the file There is no record of confirmation from the client that any payment was received 	Type whole dollar amount	Close letter Terms of settlement Court orders
AI	Closed file as lost contact with client	Yes No	Close letter

REFERENCES

Australian Bureau of Statistics. (2009). ABS data quality framework (ABS Catalogue No. 1520.0). Canberra: Australian Bureau of Statistics.

Australian Institute of Health and Welfare. (2019). Data Governance Framework 2019. Canberra: Australian Institute of Health and Welfare.

Broadhurst, K., Cusworth, L., Harwin, J., Alrouh, B., Bedston, S., Trinder, L., Jones, K., Ford, D., & Griffiths, L. (2021). Scaling up research on family justice using large-scale administrative data: an invitation to the socio-legal community. *Journal of Social Welfare and Family Law*, 43(3), 237-255.

Buhler, S., & Kaiser-Derrick, C. (2020). Home, Precarious Home: A Year of Housing Law Advocacy at a Saskatoon Legal Clinic. *Journal of Law & Social Policy*, 32, 45-66.

Byrom, N. (2019). *Digital Justice: HMCTS data strategy and delivering access to justice*. London: The Legal Education Foundation.

Connelly, R., Playford, C.J., Gayle, V. & Dibben, C. (2016). The role of administrative data in the big data revolution in social science research. *Social Science Research*, 59, 1-12.

Coumarelos, C., Macourt, D., People, J., McDonald, H. M., Wei, Z., Iriana, R., & Ramsey, S. (2012). *Legal Australia Wide Survey: Legal need in Australia*. Sydney, New South Wales: Law and Justice Foundation of New South Wales.

Curran, L. & Crockett, A. (2013). Measuring legal services: a practical methodology for measuring the quality and outcomes of legal assistance services. *University of Tasmania Law Review*, 32(1), 70–95.

Federation of Community Legal Centres. (2017). *Victorian community legal sector outcomes measurement framework*. Melbourne: Federation of Community Legal Centres.

Iwig, W., Berning, M., Marck, P. & Prell, M. (2013). *Data Quality Assessment Tool for Administrative Data*. Federal Committee on Statistical Methodology. <https://www.fcsm.gov/assets/files/docs/DataQualityAssessmentTool.pdf>.

Jones, A., Warren, I., Boxall, J., Fenn, S., Galbraith, J., Hettige, S., & Nandanakumar, M. (2022). *Understanding Civil Law Needs of Older People and Building Research Capacity in Community Legal Centres*. Melbourne: Deakin University

Jones, K.H., Heys, S., Tingay, K., Jackson, P. & Dibben, C. (2019). The good, the bad, the clunky: Improving the use of administrative data for research. *International Journal of Population Data Science*, 4(1), 1–11.

McDonald, H.M., McRae, C., Balmer, N.J., Hagland, T., & Kennedy, C. (2020). Apples, oranges and lemons: The use and utility of administrative data in the Victorian legal assistance sector. Melbourne: Victoria Law Foundation.

McLennan, D. (2018). Data quality issues in administrative data. *Administrative Data Research Network*.

Rygl, M. (2021). *Lessons from practice: Shared insights from grant funded projects*. Melbourne: Victoria Law Foundation.

Senate Economics References Committee. (2022). *Systemic, sustained and shameful: unlawful underpayment of employees' remuneration*. Canberra: Commonwealth of Australia.

Smith, M., Lix, L.M., Azimae, M., Enns, J.E., Orr, J., Hong, S., & Roos, L.L. (2018). Assessing the quality of administrative data for research: A framework from the Manitoba Centre for Health Policy. *Journal of the American Medical Informatics Association*, 25(3), 224-229.

Statistics Canada. (2017). Use of administrative data (available from <https://www150.statcan.gc.ca/n1/pub/12-539-x/2009001/administrative-administratives-eng.htm>).

Statistics New Zealand. (2016). *Guide to reporting on administrative data quality*. Statistics New Zealand.

StataCorp LLC. (2019). *Stata/SE 16.1 for Windows*. College Station, Texas: StataCorp.

Szczepanska, J., & Blomkamp, E. 2020. *Seeking Legal Help Online: Understanding the 'Missing Majority'*. Melbourne: Justice Connect.

Victoria Legal Aid. (2016). *Case study of the Below the Belt Phone App: Learning from projects that don't meet expectations* (available from <https://www.legalaid.vic.gov.au/sites/default/files/vla/vla-case-study-below-the-belt-phone-app.pdf>).

Witte, S. (2020). Case file analyses in child protection research: Review of methodological challenges and development of a framework. *Children and Youth Services Review*, 108, 104551.

Workplace Gender Equality Agency. (2017). *Data Governance Framework*. Canberra: Workplace Gender Equality Agency.



SOUTH-EAST MONASH LEGAL SERVICE INC.

E: INFO@SMLS.COM.AU

P: (03) 9545 7400 | (03) 9038 8002

W: SMLS.COM.AU