Enhancing Early Detection of Cognitive Impairment in the Criminal Justice System: Feasibility of a Proposed Method

Ruthie Jeanneret, Caroline Spiranovic, Lisa Eckstein, Rebekah McWhirter, Anna Arstein-Kerslake, Joel Scanlan, Kenneth Kirkby, Paul Watters, James Vickers

Abstract

Persons with cognitive impairment(s) are overrepresented in the criminal justice system (CJS) yet many instances of cognitive impairment go undiagnosed. As this article outlines, it would be both desirable and feasible to use automated alerts to flag accused persons who may require assistance in interacting with the CJS either due to a confirmed or likely diagnosis of a cognitive impairment or other relevant condition. A proposed method to develop this alert system is outlined, combining Natural Language Processing (NLP) applied to Electronic Health Records (EHRs) with data linkage (DL) of health and CJS data. Although there are technical barriers, this article focuses on the ethical and legal barriers of this proposed approach. It is concluded that the overall benefits of the proposed alert system would likely outweigh potential adverse outcomes. It is argued that a waiver of consent would be appropriate and that legal barriers, in terms of privacy legislation at both Federal and State levels, which apply varying requirements for the disclosure of personal information, may be overcome in part through de-identification strategies. The examples provided in this paper of criminological data linkage projects support the feasibility of the method proposed.

Word count: 7909

Key words: Criminal justice system, cognitive impairment, early detection, natural language processing, data linkage, electronic health records, ethics, privacy.

Author affiliations: Ruthie Jeanneret, Faculty of Law, College of Arts, Law and Education, University of Tasmania, Private Bag 89, Hobart, Tasmania, Australia 7001, + 61 3 6226 2066, ruthie.jeanneret@utas.edu.au Dr Caroline Spiranovic, Faculty of Law, College of Arts, Law and Education, University of Tasmania, Private Bag 89, Hobart, Tasmania, Australia 7001, + 61 3 6226 4847, caroline.spiranovic@utas.edu.au; Dr Lisa Eckstein, University of Tasmania, Faculty of Law, College of Arts, Law and Education, University of Tasmania, Private Bag 89, Hobart, Tasmania, Australia 7001, +61 3 6226 2064, lisa.eckstein@utas.edu.au; Dr Rebekah McWhirter, University of Tasmania, Centre for Law and Genetics, Faculty of Law, College of Arts, Law and Education, University of Tasmania, Private Bag 89, Hobart, Tasmania, Australia 7001, +61 3 6226 4605, Rebekah.mcwhirter@utas.edu.au; Dr Anna Arstein-Kerslake, Melbourne Law School, Mezzanine Level, GM26, 185 Pelham Street, The University of Melbourne, Victoria 3010 Australia, +61 3 8344 3806, anna.arstein@unimelb.edu.au; Dr Joel Scanlan, Discipline of ICT, School of Technology, Environments and Design, College of Sciences and Engineering, University of Tasmania, Private Bag 87 Hobart TAS 7001, +61 3 6226 7825, joel.scanlan@utas.edu.au; Professor Kenneth Kirkby, School of Medicine, University of Tasmania, Private Bag 34, Hobart TAS 7001, +61 3 6226 4802, Ken.Kirkby@utas.edu.au; Professor Paul A. Watters, Department of Computer Science and Information Technology, La Trobe University, Bundoora VIC 3068, +61 3 9479 3415, P.Watters@latrobe.edu.au; Professor James Vickers, Wicking Dementia Research and Education Centre, University of Tasmania, Private Bag 143, Hobart TAS 7001, +61 3 6226 4808, James.Vickers@utas.edu.au

Corresponding author: Caroline Spiranovic, Faculty of Law, College of Arts, Law and Education, University of Tasmania, Hobart, 7001, Australia, Email: caroline.spiranovic@utas.edu.au

Introduction

It is widely recognised that individuals with cognitive impairments, including those with mental disorders, acquired brain injuries, fetal alcohol spectrum disorder, dementia, and a range of developmental intellectual disabilities, are overrepresented in the criminal justice system (CJS). Notably, it may be that a person has an unidentified cognitive impairment upon entry into the CJS or that they acquire such an impairment during incarceration (due to for instance traumatic brain injury, alcoholism/substance abuse, or dementia). However, many instances of these cognitive impairments are undiagnosed. It is imperative that individuals with cognitive impairments who come into contact with the CJS are identified as early as possible as having an impairment that may, if recognised and appropriately accommodated, improve their opportunities for equal access to justice and participation. Cognitive impairment covers a range of neuropsychological deficits that are associated with disabilities (restricted ability to perform activities) and underlying conditions and diseases (see for instance Mental Health Commission of NSW 2017; Stewart, Wilton & Sapers 2016). These deficits and disabilities may limit a person's opportunities for equal access to and participation in justice processes (e.g. see White, Meares & Batchelor 2014). The focus of this paper is on the over-representation of people with cognitive impairments in the CJS and the ethical and legal feasibility of a proposed method to enhance early detection of such individuals.

Improved identification of cognitive impairment and the corresponding provision of appropriate services in the CJS is a human rights issue (Human Rights Watch 2018). Articles 12 and 13 of the *United Nations Convention on the Rights of Persons with Disabilities* highlight that, in order to respect the rights to equal recognition before the law and access to justice, persons with disabilities must, wherever possible, be given every support and opportunity to participate in the CJS. This includes the opportunity to exercise their legal capacity and enter a plea, as well as their ongoing participation in the trial process and other proceedings throughout their time in the CJS. In order to achieve this, there must be appropriate support provided throughout this time. When such individuals are not identified or provided with appropriate supports in interactions with the CJS, there is a risk of breaching not only these rights but also their rights to due process, more generally (Australian Human Rights and Equal Opportunity Commission et al. 1989). This is at odds with ethical and legal principles of justice and fairness, which are central tenets of the CJS.

In addition to potential ethical and human rights breaches that may occur when individuals with cognitive impairments go unrecognised and unsupported, there are additional reasons for identifying such individuals as early as possible either before or when they first come into contact with the CJS. Australian researchers in particular (e.g. see Baldry, Briggs, Goldson & Russell 2018; Baldry, Clarence, Dowse, & Trollor 2013) have highlighted the complex connections between cognitive impairment, social disadvantage, complex needs, vulnerability to harms and involvement in the CJS. The presence of one of these adversities increases the likelihood of possessing another. Furthermore, those presenting with multiple adversities are likely to have earlier and

more frequent engagement with the criminal justice system. For these and other reasons, there have been numerous calls to improve and reform the fragmented response in Australia of the CJS to people with cognitive impairments (e.g. see Mental Health Commission of NSW 2017) and also to improve screening processes to identify individuals presenting with symptoms of cognitive impairment (see Human Rights Watch 2018).

Despite significant variation in statistics on prevalence rates of cognitive impairment within the CJS, owing in part to differing definitions of (Hellenbach, Karatzias & Brown 2017), and methodologies for identifying (Bronson, Maruschak & Berzofsky 2015) cognitive impairment, various scholarly sources acknowledge higher rates of cognitive impairment within prison populations, as compared to the general population. Some statistics have placed prevalence rates of cognitive impairment as high as 69%, and some have placed it as low as 4% (Hellenbach et al. 2017; Murphy, Gardner, & Freeman 2015). In Australia, it is estimated the prevalence of intellectual disabilities, just one type of cognitive impairment, is between 8 and 15% for offender cohorts compared to 3% in the general population. This figure is higher when offenders who have a borderline IQ are included in statistics (Shepherd, Ogloff, Shea, Pfeifer, & Paradies 2017).

Individuals with an average IQ who nonetheless exhibit characteristics of cognitive impairment are not captured in these statistics (Stewart et al. 2016). A key recommendation of a 2014 Australian Law Reform Commission (ALRC) report was that the Australian common law test of 'unfitness to stand trial' sets too high a threshold and is at odds with modern aims of the trial process. Arstein-Kerslake et al (2017, p. 400) similarly stated that 'there are persons with cognitive disabilities across the globe whose rights to legal capacity and a fair trial are being breached by unfitness to plead laws.' The application of the test has also been too focussed on intellectual ability, without recognising that 'cognitive impairment resulting from any neurological condition may not be detected on an IQ test, but may still significantly impact on a person's ability to comprehend, communicate, and actively partake in legal proceedings' (White, Meares & Batchelor 2014, p. 78). Even where an individual is declared unfit to stand trial, the result may be that the individual is indefinitely detained and may remain in custody for extended periods (ALRC 2014). Vulnerable and disadvantaged populations disproportionately experience such negative outcomes (e.g. see Shepherd et al. 2017).

As eluded to earlier, incarcerated persons may also develop or acquire a brain injury post trial or during incarceration. For instance, ageing in place within the prison may expose those incarcerated to conditions such as dementia (Williams et al. 2012). Furthermore, significant risk factors for dementia, such as low educational attainment, low physical activity, depression, stress, smoking and poorly managed vascular disease may either precede or accompany incarceration (Christodoulou 2012). Hence it is important to not only identify individuals with cognitive impairments in need of assistance at the pre-trial stage but also post-trial and during incarceration.

In this paper, we assess the potential utility of a method for early identification of individuals with cognitive impairment although the proposed method could be used

at any point in the CJS to aid identification. Early detection of cognitive impairment is essential to enhancing positive outcomes such as rehabilitation, crime prevention, and fair and just treatment, as well as preventing negative outcomes including reoffending and exacerbating pre-existing social disadvantage. An appealing practical solution to enhancing early detection is for criminal justice agencies to employ routine screening for cognitive impairment (Ogloff 2015). However, it is unclear at what point such screening should be undertaken and by whom. Ideally screening would occur at the earliest point of contact with the CJS so that individuals with cognitive impairments may be diverted where possible to more appropriate and potentially therapeutic avenues.

This suggests that screening could or should be undertaken by the police, who are typically the gatekeepers to the CJS. But could police officers themselves undertake screening assessments or should skilled clinicians do this? Given the sheer numbers of accused persons entering the system and relative shortage of skilled clinicians, it would be unrealistic to have trained clinicians on hand in every police station to undertake a gold standard assessment of each suspect. It seems more realistic that police themselves could undertake this screening using a validated screening tool. However, such tools do not necessarily cover the full range of cognitive impairments and disabilities that may impede an individual's ability to participate equally in and achieve equal access to justice. Furthermore, a screening tool may still require some degree of clinical interpretation and training. Even if police could undertake such screening, another pertinent consideration is when screening should be conducted - before a suspect is interviewed or before a suspect is brought into custody? These issues aside, requiring police to undertake such screening would pose an unrealistic strain on resources both in terms of the time required to undertake the screening and in terms of the training required of police officers. Hence, an approach to early detection is required that would not place undue demands on already strained resources in the CJS while identifying individuals likely to require further assessment and possible assistance, with a relatively high degree of accuracy. Here we propose and explore the feasibility of a research method, involving Natural Language Processing (NLP) applied to Electronic Health Records (EHRs) combined with data linkage of health and CJS data, to develop automated alerts to flag individuals with a diagnosed cognitive impairment or who otherwise possess symptoms indicative of a cognitive impairment. These automated alerts could be used at multiple points throughout the CJS.

There is increasing recognition of the potential power of secondary research uses of the data contained in EHRs to transform the quality of life for the population as a whole through advances in knowledge about diseases, illnesses and impairments and how to more effectively treat or even prevent them. However, owing in part to the Value, Velocity, Volume, Veracity and Variety (the five Vs) of big data as well as privacy constraints (Watters et al. 2009), sophisticated methods of data extraction and analysis are required. We propose here the potential of a data mining method, NLP, to extract and analyse EHR data. In the context of enhancing early detection of cognitive impairment, NLP could be applied to text-based notes of clinicians contained in EHRs, extracting information identified as indicating a strong likelihood of cognitive

impairment. Information extracted from these records could form the input for classifiers and other algorithms to identify and predict patient characteristics using Machine Learning (ML). Cases identified as red flags in such EHRs could, in turn, be linked with CJS data so that CJS officials, at multiple points throughout the system, are notified early in any interactions with persons suspected of cognitive impairment. Although there are a number of technical difficulties that would need to be surmounted[1], the purpose of this review is to assess the ethical and legal feasibility of our proposed research method.

Electronic health records and their secondary uses

EHRs are 'health records of patients, which are securely stored in electronic form and shared between health-care providers' (Spiranovic, Matthews, Scanlan, & Kirkby 2016, p. 14). This definition includes 'electronic medical records, computerised patient records, computerised medical records, personal health records, and clinical data repositories' (Spiranovic et al. 2016, p. 14). A more detailed definition has been outlined in Häyrinen, Saranto, and Nykänen (2008).

Primary uses of the clinical data within EHRs are uses that are directed at the care of the individual patient. Secondary uses go beyond the individual patient, and have been defined as 'non-direct use of personal health information including but not limited to analysis, research, quality/safety measurement, public health, payment, provider certification or accreditation, and marketing and other business including strictly commercial activities' (Meystre, Lovis, Bürkle, Tognola, Budrionis, & Lehmann 2017, p. 38). Access to clinical data for secondary uses is advantageous for many reasons, including the potential to promote higher quality healthcare, improved health management, reducing costs, and broader population health management and research (Meystre et al. 2017).

Secondary use of clinical data has been facilitated by a range of sophisticated data processing methods including data mining technologies such as NLP (see Liao, Cai, Savova, et al. 2015). Prior to the availability of data mining, extracting data from clinical notes required someone to manually read and extract data from hard-copy versions and only then could data be manually entered into a system to identify patterns (Meystre et al. 2017). NLP has been transformative (Liao et al. 2015). NLP has successfully been used, among other things, to detect warning signs of heart failure, to screen and predict adverse responses to certain drugs, and for genetic research (see Spiranovic et al. 2016). Furthermore, the CRIS-CODE project more recently provides some guidance for expanding NLP into mental illness. This project aimed to use NLP 'to offer comprehensive profiling from the mental health electronic record of symptoms and of interventions, outcomes and other relevant contextual factors currently only available from text fields' (Jackson, Patel, Jayatilleke, Kolliakou, Ball, Gorrell, et al.

---

[1] In addition, NLP combined with ML would never be 100% accurate. Misclassification may result in incorrect decisions including (a) incorrectly identifying an individual as having an impairment, or (b) missing an impairment that is actually present. Hence, research in this area should seek to identify rates of false positives and false negatives.

2017, p. e012012). The results of the project demonstrated the potential for using NLP to extract symptoms of severe mental illness from patient discharge summaries, and identify common symptomology (Jackson et al. 2017).

In reviewing the applications of NLP, it seems plausible and feasible to apply NLP paired with Machine Learning (ML) techniques[2] to clinicians' notes contained in EHRs to enhance the detection of cognitive impairments. Both clinician-informed and statistically derived algorithms could be used to devise a list of terms associated with the various impairments that are included under the umbrella term cognitive impairment. EHR cases identified as likely to have a cognitive impairment and other cases in the EHR could then be compared on the actual known diagnosis[3] as entered in the EHR to establish the validity of such an approach.

Use of data linkage methods for criminological research

Although it is at least plausible, in terms of our proposed research method, to apply NLP to EHRs to enhance early detection of cognitive impairment, what of the potential for enhancing early detection in the CJS? Although enhancing early detection of cognitive impairment in healthcare settings is an important outcome in itself, this information will not aid in identifying persons with cognitive impairment in the CJS if it is not accessed by CJS agencies. Linking EHR data with CJS databases would allow a red flag to be raised for those individuals who are likely to or have a confirmed diagnosis of cognitive impairment, at the point that a suspect's or defendant's details are entered. Such linkage of existing datasets represents a novel and potentially time- and cost-effective means of addressing an urgent problem facing the CJS.

DL is essentially 'the bringing together from two or more different sources, data that relate to the same individual, family, place or event' (Holman, Bass, Rosman et al. 2008, p. 767).' There are already numerous data linkage initiatives across Australia, particularly in the health sector. One noteworthy example is the Personal Health Research Network (PHRN) which aims to provide researchers access to de-identified data to assist in population-level research (Boyd, et al. 2012). However, Ferrante (2009, p. 380) has argued that 'compared with developments in the health sector, DL-based research in the Australian criminal justice sector is in almost embryonic form'. A decade ago now, Ferrante identified a few very promising Australian examples of DL in the CJS context. More recently, Baldry, Dowse and Clarence (2012) linked NSW data from human services and criminal justice agencies for 2,731 people in prison who were known to have a cognitive impairment in order to better understand offending patterns. The study indicated that 'those with complex cognitive disability (ie comorbidity/dual diagnosis)' were more likely to have earlier and more frequent contact with the CJS, to be homeless and to have limited education (Baldry et al. 2012, p. 14). Another example of criminological DL research is the Queensland Linkage

---

[2] On a technical note, NLP extracts the data, but does not provide detection or predictive ability. Machine Learning (ML) provides the latter.

[3] The authors propose to use a forensic patient database in this initial validation work so that data on known diagnoses is reliable.

Project, which has collected longitudinal administrative data on three separate cohorts who have had contact with child protection systems as well as the youth and adult CJSs (Stewart et al. 2015). These examples demonstrate the possibility and potential benefits of conducting research using DL and applying it to CJS issues.

While these pioneering studies show the feasibility of using DL in criminological research, legal barriers may limit widespread use. State and federal legislation limits the use of personal information, including clinical notes from EHRs, which has implications for the use of information for research generally, and particularly research involving the CJS. This is because the *Privacy Act 1988 (Cth)* ss 95-95A and some state/territory privacy acts limit research exemptions to 'medical and health research', which does not include most CJS research. The issue of privacy will be explored in greater depth in the following section.

Ethical and legal issues

Secondary use of personal health data

Privacy concerns have dominated public debate about the use of EHRs and are a key challenge for the expansion of secondary uses (Caine & Hanania 2013). Under Commonwealth and state/territory privacy legislation, personal information, including electronic health information, can only be used for the purpose for which it was collected. Exceptions apply where a person has given consent or for specified other activities deemed to be in the public interest. In most jurisdictions, this includes the exceptions for health and medical research.

Relevant Australian privacy laws and research guidelines

The data relevant for early detection of cognitive impairments will predominantly comprise personal health records, potentially governed by multiple privacy regimes. The *Privacy Act 1988* (Cth) applies to records from private healthcare providers, including many general practitioners, as well as any Commonwealth-held data, such as Medicare Benefits Schedule (MBS) and Pharmaceutical Benefits Scheme (PBS) data. State/territory privacy laws will apply to health records held by public hospitals. Some state/territory laws (e.g., NSW, Vic) also apply to health records of private providers, leading to multiple layers of regulation.

Integrating early detection of cognitive impairments with the CJS adds further complexity. CJS data may be covered by Commonwealth, state/territory privacy laws. Table 1 below provides a summary of the relevant federal, state/territory privacy legislation in Australia, specifically as it relates to the use or disclosure of health and non-health related data for research and the role of the Human Research Ethics Committee (HREC). The *National Statement on Ethical Conduct in Human Research* (National Statement) is the primary set of guidelines for human research in Australia (National Health and Medical Research Council, the Australian Research Council and the Australian Vice-Chancellors' Committee 2015). The National Statement generally

requires research participants (including persons participating by virtue of use of their personal information) to provide explicit, voluntary, and informed consent. However, in some circumstances these consent requirements are waived. For a waiver of consent to be approved, the Committee must be satisfied, among other things, that the benefits from the research justify any associated risks of harm, obtaining consent is impracticable, and that there is no likely reason for thinking that participants would not have consented if they had been asked. These requirements apply both to health and non-health related data.

Table 1. Summary of privacy legislation on disclosure of health and non-health related data for research

Jurisdiction	Use or Disclosure of Health Data	Use or Disclosure of Non-Health Data
Cth	The use or disclosure of health information is necessary for research relevant to public health or public safety and it is impracticable to obtain consent. Must also comply with guidelines issued by the NHMRC, which include a requirement for HREC approval (*Privacy Act 1988* s 16B).	No clear research exception for non-health information in the *Privacy Act 1988*.
NSW	The use of the information is reasonably necessary for research in the public interest, and either it is impracticable to obtain consent or reasonable steps are taken to deidentify the information. It must also comply with guidelines issued by the Privacy Commissioner (*Health Records and Information Privacy Act 2002* sch 1, ss 10-11).	The use or disclosure is necessary for research in the public interest and either it is impracticable to obtain consent or reasonable steps are taken to deidentify the information. It must also comply with guidelines issued by the Privacy Commissioner (*Privacy and Personal Information Protection Act 1998,* s 27B).
Vic	The use or disclosure is necessary for research in the public interest, it is impracticable for the organisation to seek consent and the information is in a deidentified form (*Health Records Act 2001* sch 1, s 2).	The use or disclosure is necessary for research in the public interest and it is impracticable for the organisation to seek consent (*Privacy and Data Protection Act 2014* sch 1)

Qld	The use or disclosure is necessary for research related to public health or safety, it is impracticable to obtain consent, is conducted in accordance with guidelines approved by the chief executive, and, for disclosure, that agency reasonably believes that the information will not be disclosed further (*Information Privacy Act 2009* sch 4, s 2(c)).	The use or disclosure is necessary for research in the public interest, the information is in a deidentified form, and seeking consent is impracticable (*Information Privacy Act 2009* sch 3, s 10(f)).
WA	No legislative privacy regime	No legislative privacy regime
SA	No specific state legislative privacy regime. There does not appear to be a research exception for health related data in the Government of South Australia Administrative Instruction (2016).	No specific state legislative privacy regime. There does not appear to be a research exception for non-health related data in the Government of South Administrative Instruction (2016).
TAS	In addition to the use and disclosure exception relevant to all personal information, health information can be collected for research relevant to public health and public safety if it is in a deidentified form, it is impracticable to obtain consent, and the information is collected as required by law by competent professionals (*Personal Information Protection Act 2004* sch 1, s 10(4)).	The use or disclosure is necessary for research or statistics in the public interest and is not in a form that identifies any particular individual, and the agency reasonably believes that the recipient will not disclose the information. (*Personal Information Protection Act 2004* sch 1, s 2(c)).
ACT	The use or disclosure is necessary for research in the public interest, it is impracticable to obtain consent, and the information is deidentified (*Health Records (Privacy and Access) Act 1997,* sch 1, s 10(3)).	No specific research exceptions in the *Information Privacy Act 2014.*
NT	If use or disclosure is in the public interest, no individual will be identified, it is impracticable to obtain consent, and it is reasonably believed that the recipient will not disclose the information. Health information must also comply with any guidelines issued by the	If use or disclosure is in the public interest and is in a deidentified form, it is impracticable to obtain consent, and it is reasonably believed that the recipient will not disclose the information (*Information Act 2002* sch 2, s 2.1(ca)).

	Commissioner (*Information Act 2002* sch 2, s 2.1(ca); s 86(1)(a)(iv)).	

The requirements for release have broad similarities. Notably, in all jurisdictions with a specific legislative privacy regime, there are exceptions to use or disclosure of health data if the research is in the public interest, it is impracticable to obtain consent and/or data can be released in de-identified form. The lack of clear research exceptions for non-health data (and, in some jurisdictions, for any personal information) can prevent DL initiatives, including those that link health and non-health data such as the project seeking to detect cognitive impairments in the CJS.

The prioritisation of health and medical data in a number of Australian jurisdictions is out of step with international comparators. Under s 33(1) of the UK *Data Protection Act 1998*, 'research purposes' are defined as statistical and historical purposes, without recourse to health or medical parameters. The Act exempts data processed for research purposes from the proscription on use and disclosure if the data will not identify a specific individual, and the research activity is unlikely to cause damage or distress to persons to whom the information relates.

In Canada, the use and disclosure of personal information is regulated under s 7(2)(c) of the *Personal Information Protection and Electronic Documents Act 2000 SC 2000*. Under s 7(3)(f) of the Act, disclosure of personal information is permissible for research purposes, where the purpose cannot be achieved without such a disclosure; it is impracticable to obtain the consent of persons to whom the information relates; and the organisation informs the Commissioner of the disclosure before it occurs. The NZ *Privacy Act 1993* specifies an even more permissive regime excepting under Principle 2(g), information used for research purposes if it is de-identified and will not be published in a form where the individual is identified. Again, this applies equally to health and non-health data.

There is no clear policy reason for differentiating between health and non-health data for research, particularly given the often-blurry line between these categories. In its pivotal 2008 report into privacy laws, the ALRC recommended that the *Privacy Act* be amended to extend research exceptions for health and medical data to cover human research more generally (2008). An emerging emphasis on 'open data' in Australia and elsewhere (see for instance https://researchdata.ands.org.au) may help to shift this paradigm further in favour of criminological research in the future. Having said this, there are several factors that may justify a more cautious approach where criminological projects are concerned including the sensitive subject matter, the vulnerable and typically disadvantaged target group and the inherent power imbalance between corporate/state bodies and the individual (Israel 2004).

Addressing privacy concerns

Deidentification methods for electronic health records

In considering the relevance of privacy laws to the proposed method for enhancing detection of cognitive impairment in the CJS, it is pertinent to note that privacy laws only apply to 'personal information', which is generally limited to information about an identified or 'reasonably identifiable' individual (see for instance *Privacy Act 1988* (Cth), s 6). To the extent that research can proceed with sufficiently de-identified information, researchers will be able to circumvent most legal and ethical privacy concerns. However, by definition, DL requires information about the same individual using patient identifiers, such as a date of birth or postcode, which carries an inherent risk of re-identification. Generally, the DL process is as follows: an important life-event occurs; data is collected and stored; technology is used to link records belonging to the same person from different events; numbers or 'linkage keys' are created to identify the records as belonging to the same person; and finally, linkage keys are stored (Department of Health WA 2017). These encrypted linkage keys can sometimes be extracted and used in research (Department of Health WA 2017).

Strategies to circumvent the legal and ethical concerns are usually straightforward including the use of DL keys. Research ethics committees also commonly issue waivers of consent for DL research projects when the first 'use' of data is deidentification, provided adequate security protocols are in place (El Emam 2011). In some instances, research carries minimal risk but where the information is sensitive there is more potential for harm or discrimination as a result of reidentification or misuse of information (Porsdam Mann, Savulescu & Sahakian 2016). The risks are therefore likely heightened in our proposed method for enhancing detection of cognitive impairment in the CJS as this involves not only data linkage of sensitive health and CJS data but also the analysis of clinician entered free text in EHRs.

A particularly challenging problem with our proposed method is the potential for identifying information (e.g. names of people/locations, diagnosis, descriptions of criminal behaviours) to be retained in EHR free text. Free text—the essential element of NLP—is harder to strip of identifying information, where more specific and limited data items are being collected. Algorithms have been developed to de-identify clinical notes, which report quite high levels of accuracy (e.g. see Fernandes et al. 2013). However, some potential for privacy breaches remain, especially in instances of data entry errors or incorrectly spelled information that systems may not recognise (Fernandes et al. 2013).

Information extraction algorithms also encounter some errors due to lexical ambiguity and other constraints, which appear hard to completely address, due to overlap between personal identifiers and medical or other common use terms, such as names that also operate as physical attributes like 'Green' and 'Brown' (Morrison, Li, Lai & Hripcsak 2009). Rich text can also lead to re-identification even in the absence of traditional patient identifiers. One example of re-identification occurred using 'de-identified' medical records that included the profession of the patient's father and his

region of work (El Emam 2011). Systems are likely to improve but will not provide a complete solution. This is especially true for small jurisdictions and for rare diseases because relational information can more easily result in probabilistic identification of a patient (El Emam 2011). In any case, there is increasing recognition that true de-identification is less feasible in the age of big data/data linkage and that re-identification is always a risk (e.g. see Lipworth, Mason, Kerridge & Loannidis 2017).

Despite these difficulties in achieving deidentified data, from a public good perspective (e.g., see Ballantyne & Schaefer 2018), the overall perceived benefits of the proposed method particularly in terms of addressing inequities in access to justice for vulnerable persons, notwithstanding risks, which may be mitigated, may justify a waiver of consent.

Consent considerations

There are some unique considerations around consent raised by our proposed research. In particular, the target population is a vulnerable population. Section 4 of the National Statement (NHMRC et al. 2015) outlines the unique ethical considerations for research involving persons with a cognitive impairment. One noteworthy challenge is the likelihood that some patients will not (or will no longer) have the ability to consent to sharing health information. Nonetheless, human rights and disability rights frameworks point towards the importance of principles such as inclusion and respect for autonomy which promote the involvement of persons with cognitive impairments in research provided there are adequate safeguards to reduce potential harms (e.g., see McDonald, Conroy, & Olick 2017; McDonald & Raymaker 2013).

Individual participant consent, where an allowable exception in privacy laws exists, is the gold standard for research with identifiable and/or sensitive information. However, obtaining consent is often not practicable for DL activities, given the typically large population size, potential delay between an EHR being generated and a researcher seeking access, and participation bias. Therefore, seeking specific participant consent for a given data linkage project often will not be possible/feasible particularly in retrospective research where data has already been collected. As mentioned in the preceding section, a waiver of consent may be warranted in such instances.

However, in prospective research where data is yet to be collected, implementation of electronic health systems affords an opportunity to incorporate 'privacy aware' strategies in the initial architecture. 'Broad consent' models involve a single opportunity to give consent to all future research of a particular type, rather than on a case-by-case basis (Ploug & Holm 2015). Ideally, broad consent would be obtained at the inception of an EHR (Kosseim & Brady 2008). While this overcomes some issues with obtaining consent, a key problem with any broad consent model is that participants cannot truly provide informed consent in advance as participants would be unlikely to foresee all possible research uses of their data (Kosseim & Brady 2008).

A dynamic model, on the other hand, involves information about specific data relating to each new application being provided to participants online, along with an

option to withhold consent (Ploug & Holm 2015). It has been suggested that dynamic consent models facilitate the consent process as well as *ongoing* communication between researchers and participants (Williams et al. 2015). Despite these and other apparent benefits, dynamic models are in an early stage of development and evaluation as to their practicability.

Ploug and Holm (2015) suggest a middle path: a meta consent model, which combines broad and dynamic consent. The rationale 'is simply to let individuals choose how they wish to provide consent for future secondary research of data collected in the past or of data that will be stored in the future' (Ploug & Holm 2017, p. 2). Arguably, this could provide a flexible and practical solution, as well as appropriate respect for individual autonomy (Ploug & Holm 2017, p. 3).

For any of these strategies to be successful, they need to be addressed at the time of implementation. Once an EHR system is operational, re-consent processes will likely be considered unfeasible (Kosseim & Brady 2008). Whatever method is adopted, there is a need to ensure that any consent strategies incorporated into secondary uses of EHRs are meaningful. The mere notion of personal control cannot be taken as a panacea for privacy concerns or an imprimatur of ethical acceptability (Spriggs, Arnold, Pearce & Fry 2012).

Summary and Discussion

As this paper has outlined, the legal and human rights of persons with a cognitive impairment are breached when they do not receive the support they require to participate equally and effectively in CJS proceedings. There are a multitude of underlying neurological conditions that may give rise to a cognitive impairment which impedes on the ability of a person to communicate effectively and participate in CJS processes. To address the call for improved recognition and support of all such individuals in the CJS, screening at various points of the CJS is an intuitively appealing approach. However, as outlined in this paper, there are numerous reasons to suspect that widespread screening may not be feasible. Cost-effective and efficient alternatives to screening are desirable. For these and other reasons, an alternative method for enhancing detection of cognitive impairments in the CJS was proposed. The proposed method includes NLP combined with ML of EHRs to identify from clinician entries persons who have or are likely to have a cognitive impairment. Linking these health records with CJS data in turn would allow for red flags to be issued to CJS officers when the accused person's details are entered into a CJS database.

Although there are technical barriers to the proposed method, this article focused on the ethical and legal barriers of the proposed method. A review of federal and state/territory privacy legislation highlighted that limitations on disclosure of personal information and the absence of exceptions for research in some jurisdictions and for non-health research in others may impede the widespread implementation of the proposed approach. Nonetheless, it was suggested that data linkage keys and deidentification strategies as employed in large scale successful DL initiatives may overcome, in part, this legislative obstacle. It was also suggested that the broader nature

of exceptions to disclosure as employed in some other counties combined with moves towards open data more generally may, moving forwards, promote DL initiatives such as the one proposed here.

A further barrier identified was that of consent. Where exceptions to disclosure are permitted for non-health related research, a common requirement is that consent is granted for this disclosure. In terms of the proposed method to enhance detection of cognitive impairment, re-consent would be required for data linkage involving retrospectively collected data and prospective consent would be required for any new data collection and linkage. As outlined in this paper, there are complexities involved in obtaining informed consent from a person with a cognitive impairment. However, this should not preclude the right of such individuals to participate in research. In any case, a waiver of consent for the proposed research would likely be granted by an ethics committee primarily on the grounds that it would be impractical to obtain consent due to the large-scale and retrospective nature of the research and the proposed benefits would outweigh the risks (namely identification or reidentification. Ideally though, prospective data linkage projects may be aided by a consent model inbuilt in EHRs and other health data repositories although it is unclear which model for consent (broad, dynamic or a blend of these two) may achieve the best balance between upholding individual patient rights and facilitating research advancements, which could likely lead to enhanced population health outcomes.

In conclusion, the proposed method for enhancing detection of cognitive impairment in the CJS offers considerable promise as a means to uphold the rights of accused persons with disabilities to receive the support they require in accessing and participating equally and effectively in CJS proceedings. As evidenced by a number of successful DL criminological projects in Australia, ethical and legal issues raised by the proposed research could be addressed successfully.

Legislation

Data Protection Act 1998 (UK)
Health Records Act 2001 (Vic)
Health Records and Information Privacy Act 2002 (NSW)
Health Records (Privacy and Access) Act 1997 (ACT)
Information Act 2002 (NT)
Information Privacy Act 2014 (ACT)
Information Privacy Act 2009 (Qld)
Privacy Act 1988 (Cth)
Privacy Act 1993 (NZ)
Privacy and Data Protection Act 2014 (Vic)
Privacy and Personal Information Protection Act 1998 (NSW)
Personal Information Protection Act 2004 (Tas)
Personal Information Protection and Electronic Documents Act 2000 (Canada)

References

Arstein-Kerslake, A, Gooding, P, Andrews, L, & McSherry, B 2017, 'Human Rights and Unfitness to Plead: The Demands of the Convention on the Rights of Persons with Disabilities', *Human Rights Law Review,* vol 17, no, 3, pp. 399-419.

Australian Human Rights and Equal Opportunity Commission, Disabled Peoples' International, Australian Council for Rehabilitation of Disabled & National Council on Intellectual Disability 1989, *The rights of people with disabilities: Areas of need for increased protection*: *Discussion Paper,* Human Rights Australia, Sydney.

Australian Law Reform Commission 2008, *For Your Information: Australian Privacy Law and Practice: Report 108*, Commonwealth of Australia.

Australian Law Reform Commission 2014, *Equality, Capacity and Disability in Commonwealth Laws*: *Discussion Paper 81,* Commonwealth of Australia.

Baldry, E, Briggs, DB, Goldson, B, & Russell, S 2018, ''Cruel and unusual punishment': an inter-jurisdictional study of the criminalisation of young people with complex support needs', *Journal of Youth Studies*, vol 21, no. 5, pp. 636-652.

Baldry, E, Clarence, M, Dowse, L, & Trollor, J 2013, 'Reducing vulnerability to harm in adults with cognitive disabilities in the Australian criminal justice system', *Journal of Policy and Practice in Intellectual Disabilities,* vol 10, no. 3, pp. 222-229.

Baldry, E, Dowse, L, & Clarence, M, 2012, February, 'People with mental and cognitive disabilities: pathways into prison'. *In Background Paper for Outlaws to Inclusion Conference.* Retrieved from https://www.mhdcd.unsw.edu.au/sites/www.mhdcd.unsw.edu.au/files/u18/pdf/MHDCDbackgroundOutlaws%20Conf1.pdf.

Ballantyne, A, & Schaefer, GO 2018, 'Consent and the ethical duty to participate in health data research', *Journal of Medical Ethics*, Published Online First 22 January 2018, doi: 10.1136/medethics-2017-104550.

Boyd, J, Ferrante, A, O'Keefe, C, Bass, A, Randall, S, & Semmens, J 2012, 'Data linkage infrastructure for cross-jurisdictional health-related research in Australia', *BMC Health Services Research,* vol 12*,* doi:https://doi.org/10.1186/1472-6963-12-480.

Bronson, J, Maruschak, LM, & Berzofsky, M 2015, 'Disabilities among prison and jail inmates, 2011–12'. *US Department of Justice Bureau of Justice Statistics.* Retrieved from https://www.bjs.gov/content/pub/pdf/dpji1112.pdf.

Caine, K, & Hanania, R 2013, 'Patients want granular privacy control over health information in electronic medical records*', Journal of the American Medical Informatics Association,* vol 20, no. 1*,* pp. 7-15.

Christodoulou, M 2012, 'Locked up and at risk of dementia', *The Lancet Neurology,* vol 11, no. 9, pp. 750-751.

Department of Health WA 2017, *Linkage and Extraction Processes*. Retrieved from http://www.datalinkage-wa.org.au/dlb-linkage-extraction-process.

El Emam, K, 2011, 'Methods for the de-identification of electronic health records for genomic research', *Genome Medicine,* vol 3, no. 25, doi:https://doi.org/10.1186/gm239.

Fernandes, A, Cloete, D, Broadbent, M, Hayes, R, Chang, C-K, Jackson, R … & Callard, F 2013, 'Development and evaluation of a de-identification procedure for a case register sourced from mental health electronic records', *BMC Medical Informatics and Decision Making* vol 13, no. 71, doi:https://doi.org/10.1186/1472-6947-13-71.

Ferrante, A 2009, The Use of Data-Linkage Methods in Criminal Justice Research: A Commentary on Progress, Problems and Future Possibilities, *Current Issues in Criminal Justice,* vol 20, no. 3, pp. 378-392.

Government of South Australia, 'PC012 - Information Privacy Principles (Ipps) Instruction' (Administrative Instructions, Department of Premier and Cabinet Circular, 20 June 2016). Retrieved from https://www.archives.sa.gov.au/sites/default/files/20160719%20Prem%20Cab%20Circ%2012%20-%20amended%20June%202016%20-%20with%20Proclamation%20FINAL.pdf.

Hawys, W, Spencer, K, Sanders, C, Lund, D, Whitley, E, Kaye, J, & Dixon, W 2015, Dynamic Consent: A Possible Solution to Improve Patient Confidence and Trust in How Electronic Patient Records Are Used in Medical Research, *JMIR Medical Informatics,* vol 3, no. 1, doi:10.2196/medinform.3525.

Häyrinen, K, Saranto, K, & Nykänen, P 2008, 'Definition, structure, content, use and impacts of electronic health records: a review of the research literature', *International journal of medical informatics,* vol 77, no. 5, pp. 291-304.

Hellenbach, M, Karatzias, T, & Brown, M 2017, 'Intellectual disabilities among prisoners: prevalence and mental and physical health comorbidities', *Journal of Applied Research in Intellectual Disabilities,* vol 30, no. 2, pp. 230-241.

Holman, CDJ, Bass, JA, Rosman, DL, Smith, MB, Semmens, JB, Glasson, EJ ... & de Klerk, NH 2008, 'A decade of data linkage in Western Australia: strategic design, applications and benefits of the WA data linkage system', *Australian Health Review,* vol 32, no. 4, pp. 766-777.

Human Rights Watch 2018 *"I Needed Help, Instead I Was Punished": Abuse and Neglect of Prisoners with Disabilities in Australia.* Human Rights Watch, United States of America. Retrieved from https://www.hrw.org/publications.

Israel, M 2004, *Ethics and the Governance of Criminological Research in Australia,* NSW Bureau of Crime Statistics and Research, Sydney.

Jackson, RG, Patel, R, Jayatilleke, N, Kolliakou, A, Ball, M, Gorrell, G ... & Stewart, R 2017, 'Natural language processing to extract symptoms of severe mental illness from clinical text: The Clinical Record Interactive Search Comprehensive Data Extraction (CRIS-CODE) project', *BMJ open,* vol 7, no.1, doi:10.1136/bmjopen-2016-012012.

Kohane, IS 2013, 'Secondary use of health information: are we asking the right question?', *JAMA internal medicine,* vol 173, no. 19, pp. 806-1807.

Kosseim, P, & Brady, M 2008, 'Policy by Procrastination: Secondary Use of Electronic Health Records for Health Research Purposes', *McGill Journal of Law and Health,* vol *2,* pp. 5-45.

Liao, KP, Cai, T, Savova, GK, Murphy, SN, Karlson, EW, Ananthakrishnan, AN ... & Churchill, S 2015, 'Development of phenotype algorithms using electronic medical records and incorporating natural language processing', *BMJ,* vol 350, doi: 10.1136/bmj.h1885.

Lipworth, W, Mason, PH, Kerridge, I, & Ioannidis, JP 2017, 'Ethics and epistemology in big data research', *Journal of bioethical inquiry,* vol *14*, no. 4, pp. 489-500.

McDonald, KE, Conroy, NE, Olick, RS, & The Project Ethics Expert Panel 2017, 'What's the harm? Harms in research with adults with intellectual disability', *American journal on intellectual and developmental disabilities,* vol 122, no. 1, pp. 78-92.

McDonald, KE, & Raymaker, DM 2013, 'Paradigm shifts in disability and health: Toward more ethical public health research', *American journal of public health,* vol 103, no. 12, pp. 2165-2173.

Mental Health Commission of NSW 2017, *Towards a just system: mental illness and cognitive impairment in the criminal justice system.* Mental Health Commission of NSW, Sydney. Retrieved from https://nswmentalhealthcommission.com.au/sites/default/files/documents/justice_paper_final_web.pdf.

Meystre, SM, Lovis, C, Bürkle, T, Tognola, G, Budrionis, A, & Lehmann, CU 2017, 'Clinical data reuse or secondary use: current status and potential future progress', *Yearbook of medical informatics,* vol 26, no. 1, pp. 38-52.

Morrison, F, Li, L, Lai, A, & Hripcsak, G 2009, 'Repurposing the Clinical Record: Can an Existing Natural Language Processing System De-identify Clinical

Notes?', *Journal of the American Medical Informatics Association*, vol 16, pp. 37-39.

Murphy, GH, Gardner, J, & Freeman, MJ 2017, 'Screening prisoners for intellectual disabilities in three English prisons', *Journal of Applied Research in Intellectual Disabilities,* vol 30, no. 1, pp. 198-204.

Ogloff, J 2015, 'Good mental health care in prisons must begin and end in the community', *The Conversation,* vol 24. Retrieved from http://theconversation.com/good-mental-health-care-in-prisons-must-begin-and-end-in-the-community-40011.

Ploug, T, Holm, S 2015, 'Meta consent: a flexible and autonomous way of obtaining informed consent for secondary research', *BMJ* vol 350, doi: https://doi.org/10.1136/bmj.h2146.

Porsdamm Mann, S, Savulescu, J, Sahakian, B 2016, 'Facilitating the ethical use of health data for the benefit of society: electronic health records, consent and the duty of easy rescue', *Philosophical Transactions of the Royal Society A,* vol 374, pp. 1-17. doi: 10.1098/rsta.2016.0130.

Shepherd, SM, Ogloff, JR, Shea, D, Pfeifer, JE, & Paradies, Y 2017, 'Aboriginal prisoners and cognitive impairment: the impact of dual disadvantage on Social and Emotional Wellbeing', *Journal of Intellectual Disability Research,* vol 61, no. 4, pp. 385-397.

Spiranovic, C, Matthews, A, Scanlan, J, & Kirkby, KC 2016, 'Increasing knowledge of mental illness through secondary research of electronic health records: opportunities and challenges', *Advances in Mental Health,* vol 14, no. 1, pp. 14-25.

Spriggs, M, Arnold, MV, Pearce, CM, & Fry, C 2012, 'Ethical questions must be considered for electronic health records', *Journal of Medical Ethics,* vol 38, no. 9, pp. 535-539.

Stewart, A, Dennison, S, Allard, T, Thompson, C, Broidy, L, Chrzanowski, A 2015, 'Administrative data linkage as a tool for development and life-course criminology: The Queensland Linkage Project', *Australian and New Zealand Journal of Criminology,* vol. 48*,* pp. 409-428.

Stewart, LA, Wilton, G, & Sapers, J 2016, 'Offenders with cognitive deficits in a Canadian prison population: prevalence, profile, and outcomes', *International journal of law and psychiatry,* vol 44, pp. 7-14.

The National Health and Medical Research Council, the Australian Research Council and the Australian Vice-Chancellors' Committee 2015, *National Statement on Ethical Conduct in Human Research 2007 (Updated 2015).* The National Health and Medical Research Council, Canberra.

Watters, PA, Kuh, D, Latham, S, Shah, I & Garwood, K 2009, Enabling access to British birth cohorts: A secure web interface for the NSHD (SWIFT). *Proceedings of the 11ᵗʰ International Conference on e-Health Networking, Applications and Services.*

White, A, Meares, S, & Batchelor, J 2014, 'The role of cognition in fitness to stand trial: a systematic review', *Journal of Forensic Psychiatry & Psychology,* vol. 25, pp. 77-99.

Williams, BA, Stern, MF, Mellow, J, Safer, M, & Greifinger, RB 2012, 'Aging in correctional custody: setting a policy agenda for older prisoner health care', *American journal of public health,* vol 102, no. 8, pp. 1475-1481.