

Australian general practitioners' attitudes to the extraction of research data from electronic health records

Adam Jose Hodgkins,

Judy Mullan, Darren John Mayne,
Costa Steven Boyages, Andrew Bonney

Background and objective

General practice electronic health records (EHRs) are a rich source of primary care data that can be used for important research. The aim of this qualitative study was to analyse the attitudes of Australian general practitioners (GPs) to the use of data extracted from primary care EHRs for clinical research.

Methods

Semi-structured interviews were conducted with 13 Australian GPs. Interviews were recorded, transcribed and thematically analysed.

Results

Two main themes emerged. The data custodian role encompassed GPs' determination to maintain privacy, their trust of organisations involved and benefits of research outcomes. The theme of protecting practice resources included concerns about unrecompensed staff time and potential risks to data and systems.

Discussion

This study highlights that while Australian GPs have concerns about the use of data stored on their EHRs, they also recognise the benefits of using this data for research purposes. Addressing these concerns could help to ensure that researchers have access to this valuable and rich data resource.

THE ELECTRONIC HEALTH RECORD (EHR) is widely used in primary care and, given the substantial amount of data stored within these systems, is a valuable source of data for research.¹ EHR data are recorded contemporaneously within the environment in which patients' health conditions and treatments occur, thus providing an important contextual factor to research.² Real-world data recorded in primary care, rather than under experimental conditions, is important for comparative effectiveness research and pragmatic trials.^{3,4} The time and cost of data collection can be reduced by using EHR data rather than other data collection methods.^{3,5}

The use of general practice EHRs as a research data source is expanding.⁵ There are a large number of primary care EHR data collection projects and substantial research output from Australia, the USA, Canada and several European countries.⁶ The UK leads the way, with over 40% of this nation's research output using primary care EHR data.⁶

However, researchers using EHR data have faced various obstacles. For example, the rollout of the Clinical Practice Research Datalink in the UK was delayed to address concerns of the medical community.^{7,8} British research highlights that a significant concern for some general practitioners (GPs) is the extraction of potentially identifiable data

from the EHR. Concerns arose that data may fall into the hands of insurers or groups other than the intended research group.^{8,9} Gaining support from colleagues within the practice was important to GPs' decisions to allow researchers access to EHR data, as was consideration of the workload that participation in research would involve for practice staff.^{8,10} Strong support from government agencies and academic institutions forged academic partnerships and was integral to the success of projects using EHR data.⁶ Overall, GPs tended to have altruistic attitudes towards using EHR data for research, but concerns about potentially identifiable data and access to data by commercial agencies were important issues. Therefore, the right to privacy must balance with genuine commitment and support of medical research.⁸

Researchers wishing to use the vast medical data stored in Australian primary care EHRs need to understand the attitudes of GPs, which are integral to their role as gatekeepers of this data.⁶ GPs' attitudes to the use of EHR data in both large scale and local projects have been well studied in the UK, Europe and USA,^{6,8-12} but appear to be limited in an Australian context. This study aims to explore Australian GPs' attitudes relating to EHR use for research, thereby informing future researchers who wish to reduce potential barriers and maximise enabling factors for GP participation.

Methods

Participants

Fifteen Australian GPs were contacted and consented to participate in semi-structured recorded interviews. No participants refused or withdrew their agreement to participate. Thirteen interviews were performed, at which point it was determined that data saturation had been reached.

Purposive sampling was used to obtain representation of GPs from different areas, of different sexes and ages, and with varying practice experience. The researchers obtained publicly available general practice contact details using Google internet searches. Practices were contacted directly by telephone, and participant information sheets and consent forms were forwarded by email for completion before interviews were arranged with the GPs.

Interviews

The primary investigator (AH), a male GP with 23 years of general practice experience, conducted all interviews. AH has an interest in using EHR data for research and has published research in this area.^{5,13} These details were disclosed to participants prior to obtaining consent.

The interviews were conducted by telephone with audio recording. A script was used to ensure coverage of major points, but free-flowing conversation with natural progression was encouraged. Interviews were conducted between May and September 2018.

Coding and analysis

Data coding was performed manually by two authors (AH and CB) using thematic analysis.¹⁴ This involved initial familiarisation of the data with reading, rereading and taking notes before progressing to generation of initial codes.

The codes were gathered into themes after discussion and input from the whole research team before assessing them against the extracts and the data as a whole. Clear definitions and identification of the core themes allowed further analysis. The consolidated criteria for reporting qualitative research (COREQ) guidelines were used to ensure rigour in the analyses and presentation of findings.¹⁵

Ethics

The research was approved by the University of Wollongong and Illawarra and Shoalhaven Local Health District Health and Medical Human Research Ethics Committee (approval number 2018/253).

Results

Thirteen GPs participated in the interviews, which ranged from 13 to 33 minutes' recorded duration. Data saturation was achieved by the eleventh interview. The characteristics of the participants, which included a broad range of ages, settings, practice sizes and years of experience, are summarised in Table 1.

Two core themes were identified through the structured analysis: 'the data custodian role' and 'protection of practice resources'. These emerged from the subthemes described below.

Data custodian role

The GP's role as custodian of the data emerged as a major theme. There was an expectation that the data use would not compromise patient privacy. GPs needed to trust the organisations involved in research that used their data. They also considered the project outcomes resulting from their data usage when appraising research participation.

Privacy

De-identification of data and privacy was a consistent subtheme identified in GPs' attitudes to the use of practice data for research. De-identification of data was viewed as essential by all participants.

I think we are always just cautious about patient privacy. (GP02)

I think you've got to have caution just about privacy and use of the data, making sure that it's quite robust in that regard to the de-identification. (GP04)

[I'm] assuming that the patient data is de-identified and that there's no risk of privacy breaches for our patients. (GP13)

The possibility of re-identification of data was considered by GPs, with concerns expressed that this may threaten privacy.

Table 1. Patient characteristics

| Characteristic | n (%) |
|------------------------------|-------------|
| Sex | |
| Female | 7 (54%) |
| Male | 6 (46%) |
| Age group (years) | |
| 40–44 | 2 (15%) |
| 45–49 | 2 (15%) |
| 50–54 | 3 (23%) |
| 55–59 | 3 (23%) |
| ≥60 | 3 (23%) |
| Rurality* | |
| MMM1 | 7 (54%) |
| MMM2 | 1 (8%) |
| MMM3 | 2 (15%) |
| MMM4 | 1 (8%) |
| MMM5 | 2 (15%) |
| State/Territory | |
| New South Wales | 8 (62%) |
| Queensland | 1 (8%) |
| Victoria | 1 (8%) |
| Western Australia | 2 (15%) |
| Australian Capital Territory | 1 (8%) |
| Practice size | |
| Solo | 1 (8%) |
| 2–5 GPs | 5 (38%) |
| 6–10 GPs | 3 (23%) |
| 11–15 GPs | 3 (23%) |
| >15 GPs | 1 (8%) |
| Owner/Contractor | |
| Owner | 9 (69%) |
| Contractor | 2 (15%) |
| Previous owner | 2 (15%) |
| Years practising† | |
| Range | 14–37 years |
| Median | 25 years |
| Mean | 26 years |

*Rurality is grouped by the Modified Monash Model (MMM) score with MMM1 being metropolitan and higher numbers representing greater remoteness.

†Years practising is reported as range, median and mean. All other characteristics are reported as n (%).

GP, general practitioner; MMM, Modified Monash Model

I guess I'd have some concerns that if it could be re-identified that there might be breaches in confidentiality. (GP05)

Re-identifiable is fine, as long I'm the only one that can ... re-identify it. (GP09)

Some participants were concerned that re-identification may be possible if rare conditions were being studied in small geographic regions.

[I]f you were doing a study on [a very uncommon treatment] then you could work out the patients' [identities]. (GP08)

Data linkage was also seen as a possible risk of re-identification. However, there was acceptance by several GPs that this was a valuable method of data collection.

[B]eing able to link up the data from the general practices with the hospital data, so you can actually start to understand where patients are going and hospital admissions. I actually think that's a really important thing to be able to do. (GP04)

The provision of aggregate data or data analysed in the practice was seen as a safe alternative to providing de-identified patient data to researchers.

I've always been a bit happier with the idea of stuff staying within the practice. (GP02)

There was recognition that ultimately the practice owners were the custodians of the data, but a number of participants stated that they would involve all doctors in the practice in a decision to participate in this type of research.

I would consult. So, it's the kind of thing we'd bring up at the practice meeting just to say there's a project going on, the project would involve tapping into our data. (GP04)

Even though one owner owns the practice, I suspect it would be a collaborative decision between all of us. I guess ultimately it's his decision. (GP12)

[T]he three practice owners will review all research projects that come in and make

the first pass. Then we'll run it out to our doctors meeting. (GP10)

Trust

Trust was an important subtheme. University-backed research projects led by GPs tended to be trusted. The majority of GPs were less trusting of the involvement of pharmaceutical and insurance organisations.

I think it's probably a little bit more about university and the team and the support – and the trust. (GP02)

Obviously, if it's a GP as opposed to a drug company, that might influence my likelihood of being involved as well. (GP12)

I think it would need to be a medical researcher ... from a tertiary institution or maybe a general practice organisation ... I don't think that we would be happy to be dealing with pharmaceutical companies and the like. (GP07)

Some GPs expressed their unwillingness to provide data to these organisations, but others decried the stereotype of large pharmaceutical corporations being inherently untrustworthy, instead pointing to the need to put outcomes ahead of the organisation in decision making.

I don't think that just because they come from the government versus come from a pharmaceutical company versus come from a university, intrinsically, as far as I'm concerned, makes them ethically any better or philosophically better or worse than anybody else. (GP10)

GPs objected to sharing data with organisations that might use it for financial gain or private interests rather than for producing outcomes beneficial for practices or patients. A GP's past experience with an organisation influenced their trust of the organisation, with anecdotes of poor past experience used to justify their current wariness.

If it's academic, if it's universities, if it's colleges ... because those organisations

don't really have anything to gain from selling my information. (GP04)

The knowledge that a research project had been through an ethics approval process was more acceptable for GPs. Even the presence of continuing professional development (CPD) points associated with research gave GPs assurance that someone other than the researcher had evaluated the project and found merit in the process.

I'd be looking at the ethics approval ... (GP07)

[S]o long as it had been before an ethics committee and that the patient's privacy was protected, I wouldn't have a problem with it. (GP13)

[CPD points] makes me think that the college has looked at the study and think that it is actually meaningful. (GP08)

Outcomes

The participants described a desire to have transparency of the data extraction, the findings and how the outcomes of the research using their EHR data might be applied.

[I]f the outcomes were [to be] published somewhere and the data will be made available to people who wanted to independently analyse results. (GP08)

So as long as the research project itself is transparent in its funding and its purposes ... (GP04)

This transparency enhanced GPs' trust, and links to the third subtheme of outcomes. GPs wanted assurance that quality data were being collected, as they were cognisant that this affected outcomes.

[M]aking sure the data that we share is actually good data [because] you know garbage in is garbage out. (GP10)

Generally, GPs were more likely to participate in research if the outcomes benefitted their patients or practices.

The topics we tend to be more interested in are the ones that are GP focused, the ones that we can see clinical value in and the sorts of questions we perhaps have been asking ourselves or that might have an impact on the physician, on the general practice, on the overall health environment, or specific things we can do for our patients. (GP02)

[B]ecause it [research] gives you ... much better pictures of how we can improve health for our community. (GP10)

Overall, there was a general altruism expressed, with participants willing to contribute to the greater good, not just advancement of their own patients' needs.

I do want to help. I mean, we go into medicine usually because we're quite altruistic. (GP06)

I'm happy ... for a researcher in the practice for the right reasons; it's an investment that's worth making. (GP04)

Protection of practice resources

While trust, outcomes and patient privacy were important for GPs, they also nominated the impact of the research on the practice as an important consideration. Time involvement was seen as a barrier; however, rewards could offset this or be seen as an enabler to involvement. Some GPs were influenced by the risk of a breach of information technology security.

Time

Time involvement was a significant subtheme cited by all participants. Clinician time spent in research was seen to detract from both income-producing work and patients' demands for attention.

It's the time-consuming nature of it, which can be a barrier ... I'm already sort of time poor, so it's just yet another thing to do. (GP01)

The idea's good but time is the constraint. (GP11)

[T]he time involvement - I'm, like most general practitioners, fairly busy. (GP07)

There was also a recognition of the importance of staff time in research participation. Many identified practice staff as the most likely to extract data. Although time was seen as a liability and barrier to research involvement, it was one that participants felt could be balanced with appropriate reward.

[W]e are a little bit cautious sometimes at the balance of the workload on the staff. (GP02)

I think if it's going to take time for admin staff to do that work, then I definitely think it would be appropriate to reimburse them sufficient to cover that. (GP12)

Reward

The subtheme of reward emerged with all participants. Remuneration was appreciated by GPs and was an enabler to participation.

If there was a monetary value attached to it, that would probably make it more appealing to me. (GP01)

However, GPs were concerned that remuneration in excess of compensation of time and effort would be problematic. Remuneration was seen by many to balance the barrier of time involvement.

Certainly, if it comes with an appropriate financial remuneration to cover some of the time ... I guess it's more about financial remuneration of time. (GP05)

[R]eimbursement for time would be really ideal. (GP10)

[J]ust ... a token amount to cover some time spent. (GP07)

[I]f it's something that it looks like I can make a profit doing, then I then worry that the data would be skewed by only having people who wanted to do this as a money making [exercise]. (GP08)

Reward did not have to be solely financial, with CPD points often considered an enabler of research participation. While

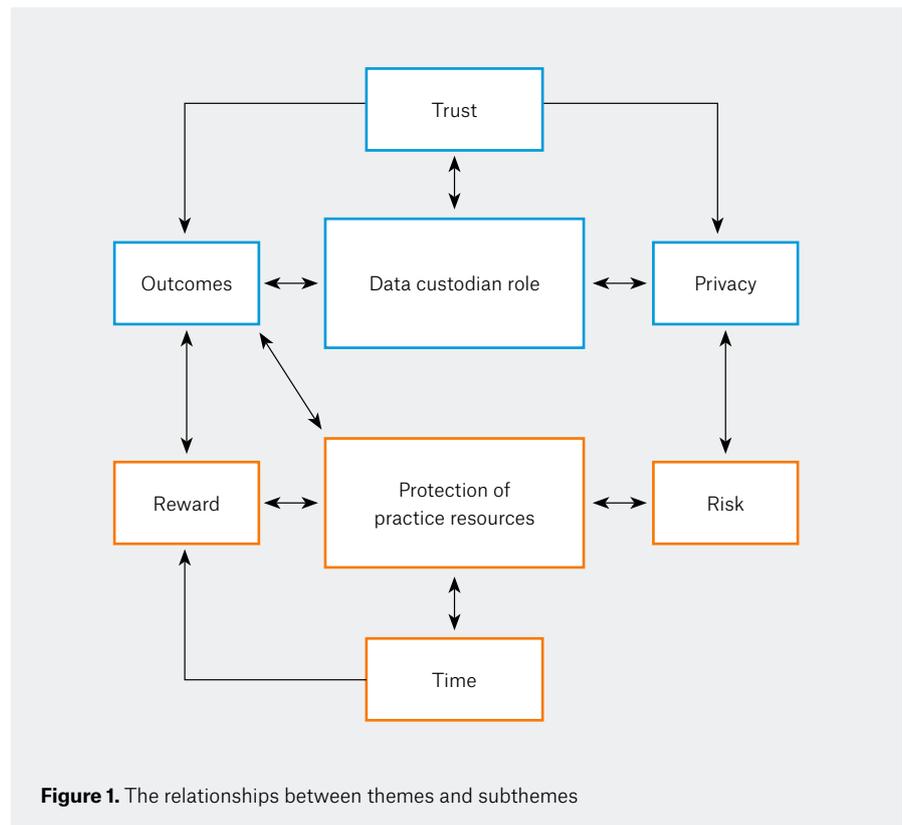


Figure 1. The relationships between themes and subthemes

some participants valued CPD more than others, no participants had negative connotations, and the points often enabled greater trust.

I think that's a nice in-kind kind of reward for involvement. (GP04)

I think practices can get a bit tired or overloaded with lots of requests, so the ones that have some kind of combination of value to the practice and their patients, CPD and some kind of financial recognition are the ones that are more likely to ... get done. (GP02)

Risk

Rewards were balanced with the final subtheme of risk. Several participants were wary of the risk of data breaches and the legal and ethical ramifications of such a breach for their practice. A systems breach was also considered a risk, with some GPs concerned about the risk of data extraction tools increasing the risk of malware affecting their system.

I would be concerned that systems might break down and ... personal information might be breached. (GP05)

I'm not prepared to put the business at risk. (GP06)

Figure 1 is a graphical representation of how the subthemes are linked and how they contribute to the construction of the main two core themes.

Discussion

The GPs that participated in this study are generally positively disposed to research, and they recognise significant advantages in drawing the data from primary care. A custodian can be defined as a 'person who has responsibility for taking care of or protecting something'.¹⁶ As custodians of the data in general practice EHRs, GPs have adopted a gatekeeper role, with the protection of patient privacy a key consideration. Greater trust in organisations reassured them of the integrity of projects, and transparency of research processes and outcomes further added trust.

The efficient running of a general practice and protection of resources such as time are important factors influencing the decision to participate in research. Remuneration can offset the time barrier and be an enabler in its own right. Avoiding risks (including data breaches) was important.

Australian research analysing the attitudes of practice staff for a specific data extraction project discussed time as a barrier and the desirability of remuneration, which supports the results of the current study.¹⁷ The present research aligns with findings from outside Australia and confirms the gatekeeper role that GPs maintain as custodians of valuable primary care EHR data.^{6,8-12} International studies discuss GPs' strong desire to protect privacy if sharing data and support the current study findings about a lack of trust in the researcher or fears that data might be used by organisations such as insurers and pharmaceutical industries.^{10,11} The barrier of time involvement and consideration of remuneration were themes in other research that concur with the present Australian study.¹⁰

Recent literature proposing best-practice principles and models for using primary care data was unable to take into account Australian GPs' attitudes, mentioning the hitherto paucity of Australian evidence.¹⁸ The results of this study support their proposed principles of data de-identification, transparency and appropriate consideration of outcomes.

These results will support researchers' ability to address attitudes of GPs in recruitment strategies for studies using EHR data. Knowledge of GPs' attitudes assists researchers to create better data extraction methods, address barriers such as time involvement and balance these issues with appropriate reward. Addressing privacy and trust are important for GPs considering involvement in primary care research using EHR data.

As well as assisting researchers to better address the attitudes of the data custodians, this research allows GPs to more clearly evaluate requests to share their EHR data for research purposes. The themes and subthemes can be used

as a framework for GPs to assess whether important aspects have been addressed. If so, they can feel confident in participating in research using the rich resource of primary care EHR data.

A possible limitation to this study is that the author performing the interviews (AH) declared an interest in using EHR data for research. This potentially could influence responses. A strength of the study is the diversity of participants with regard to practice size, rurality, state and age.

These findings describe, for the first time, the attitudes of Australian GPs towards the use of their EHR data for research. This new knowledge can potentially influence recruitment and project design and offer GPs a framework that can be used to assess requests to participate in primary care EHR collection.

Implications for general practice

- GPs' attitudes are integral to their gatekeeper role as custodians of EHR data.
- Protecting patient confidentiality and minimising disruption to practice resources are key considerations for GPs contemplating sharing patient data for research.
- Researchers wishing to use EHR data can optimise enablers and address barriers in future project design.
- GPs will be able to evaluate potential participation in EHR-based research studies with greater confidence and identify whether study designs address common GP considerations.

Authors

Adam Jose Hodgkins BMed, DipPaed, FRACGP, Clinical Senior Lecturer, Graduate Medicine, Faculty of Science, Medicine and Health, University of Wollongong, NSW. adam@hodgkins.com.au

Judy Mullan BA, BPharm, FSHPA, PhD, Associate Professor, Graduate Medicine, Faculty of Science, Medicine and Health, University of Wollongong, NSW; Director, Centre for Health Research Illawarra Shoalhaven Population, Australian Health Services Research Institute, University of Wollongong, NSW
Darren John Mayne BA (Hons), MPH (Hons), Public Health Epidemiologist, Illawarra Shoalhaven Local Health District, Public Health, Wollongong NSW; Honorary Fellow, Graduate Medicine, Faculty of Science, Medicine and Health, University of Wollongong, NSW; Affiliated Researcher, Illawarra Health and Medical Research Institute, University of Wollongong, NSW; Doctoral Candidate, University of Sydney, Sydney School of Public Health, NSW

Costa Steven Boyages MBBS, BArts, LLB, GDLP, FRACGP, D.CH, GAICD, Academic Tutor, Graduate Medicine, Faculty of Science, Medicine and Health, University of Wollongong, NSW

Andrew Bonney MBBS, MFM (Clin), PhD, FRACGP, Roberta Williams Chair of General Practice, Graduate Medicine, Faculty of Science, Medicine and Health, University of Wollongong, NSW; Affiliated Researcher, Illawarra Health and Medical Research Institute, University of Wollongong, NSW

Competing interests: None.

Funding: This research has been conducted with the support of the Australian Government Research Training Program Scholarship.

Provenance and peer review: Not commissioned, externally peer reviewed.

References

- Muller S. Electronic medical records: The way forward for primary care research? *Fam Pract* 2014;31(2):127-29. doi: 10.1093/fampra/cmu009.
- Stange KC, Miller WL, McWhinney I. Developing the knowledge base of family practice. *Fam Med* 2001;33(4):286-97.
- Cowie MR, Blomster JI, Curtis LH, et al. Electronic health records to facilitate clinical research. *Clin Res Cardiol* 2017;106(1):1-9. doi: 10.1007/s00392-016-1025-6.
- Roland M, Torgerson DJ. Understanding controlled trials: What are pragmatic trials? *BMJ* 1998;316:285. doi: 10.1136/BMJ.316.7127.285.
- Hodgkins AJ, Bonney A, Mullan J, Mayne DJ, Barnett S. Survival analysis using primary care electronic health record data: A systematic review of the literature. *Heal Inf Manag* 2018;47(1):6-16. doi: 10.1177/1833358316687090.
- Gentil ML, Cuggia M, Fiquet L, et al. Factors influencing the development of primary care data collection projects from electronic health records: A systematic review of the literature. *BMC Med Inform Decis Mak* 2017;17(1):139. doi: 10.1186/s12911-017-0538-x.
- Donnelly L. NHS medical records database halted amid concerns. *The Telegraph*. 18 Feb 2014.
- Stevenson F. The use of electronic patient records for medical research: Conflicts and contradictions. *BMC Health Serv Res* 2015;15:124. doi: 10.1186/s12913-015-0783-6.
- Stevenson F, Lloyd N, Harrington L, Wallace P. Use of electronic patient records for research: Views of patients and staff in general practice. *Fam Pract* 2013;30(2):227-32. doi: 10.1093/fampra/cms069.
- Hummers-Pradier E, Scheidt-Nave C, Martin H, Heinemann S, Kochen MM, Himmel W. Simply no time? Barriers to GPs' participation in primary health care research. *Fam Pract* 2008;25(2):105-12. doi: 10.1093/fampra/cmn015.
- Spencer K, Sanders C, Whitley EA, Lund D, Kaye J, Dixon WG. Patient perspectives on sharing anonymized personal health data using a digital system for dynamic consent and research feedback: A qualitative study. *J Med Internet Res* 2016;18(4):e66. doi: 10.2196/jmir.5011.
- Wyatt D, Cook J, McKevitt C. Perceptions of the uses of routine general practice data beyond individual care in England: A qualitative study. *BMJ Open* 2018;8(1):e019378. doi: 10.1136/bmjopen-2017-019378.
- Barnett S, Henderson J, Hodgkins A, et al. A valuable approach to the use of electronic medical data in primary care research: Panning for gold. *Heal Inf Manag* 2017;46(2):51-57. doi: 10.1177/1833358316669888.
- Braun V, Clarke V. Using thematic analysis in psychology. *Qual Res Psychol* 2006;3(2):77-101. doi: 10.1191/1478088706qp063oa.
- Tong A, Sainsbury P, Craig J. Consolidated criteria for reporting qualitative research (COREQ): A 32-item checklist for interviews and focus groups. *Int J Qual Health Care* 2007;19(6):349-57. doi: 10.1093/intqhc/mzm042.
- Lexico.com. Custodian. USA: Lexico.com, 2019. Available at www.lexico.com/en/definition/custodian [Accessed 8 November 2019].
- Ghosh A, McCarthy S, Halcomb E. Perceptions of primary care staff on a regional data quality intervention in Australian general practice: A qualitative study. *BMC Fam Pract* 2016;17:50. doi: 10.1186/s12875-016-0445-8.
- Canaway R, Boyle DI, Manski-Nankervis JE, et al. Gathering data for decisions: Best practice use of primary care electronic records for research. *Med J Aust* 2019;210 Suppl 6:S12-6. doi: 10.5694/mja2.50026.

correspondence ajgp@racgp.org.au