

#### INTRODUCTION

When Sir James Mackenzie was moving back to Scotland from London in 1918 to set up the Institute for Clinical Research in St Andrews he said that he wanted 'To do for medicine what the Atomic Theory had done for chemistry'.<sup>1</sup> He was referring to the fact that Sir Ernest Rutherford had 'split the atom' the year before and ushered in a new scientific age through nuclear fission. Mackenzie saw an analogy with what he hoped to achieve in medicine through fusion of data across time and from various sources. He believed that the information recorded in GP consultations about the early stages of disease would lead to new ways to understand health and disease, leading to earlier and better diagnosis and treatment.<sup>2</sup> He was frustrated then, as we often still are, by not knowing what the symptoms and signs our patients consult us about really mean. Mackenzie hoped that a group of family physicians who kept continuous records of all cases they saw, including details of the patients' living conditions, dietary and smoking habits, and employment, would enable research to be carried out into the relationship between environmental factors and disease.<sup>3</sup> In this idea, he was ahead of his time, but developments in using electronic medical records and practice-based research networks are now enabling his vision to be realised. This lecture will describe the extent to which the data recorded by GPs are already contributing to improved care for patients and discuss emergent opportunities in research internationally by describing how family physicians can now make significant contributions to medical science every day in their work.<sup>4</sup>

#### THE CONSULTATION AS ATOM

In a GP consultation such as that shown in Box 1 there are many data points that

#### Box 1. Data in a consultation

Robert McGregor is a 49-year-old recently unemployed accountant who made an appointment because of low mood, poor appetite, and sleep disturbance. He has non-insulin-dependent diabetes. He is also drinking two bottles of whisky per week, though he does not volunteer this information initially. You check his blood pressure, discover that it is 148/104 mmHg, and ask him to return to the practice nurse for follow-up.

*"Sir Ernest Rutherford had 'split the atom' the year before and ushered in a new scientific age through nuclear fission. Mackenzie saw an analogy with what he hoped to achieve in medicine."*

are already present in the patient's record and many others that will be added during and after the visit for his complex range of physical and psychological problems. Compare that with 1918 when the UK had only recently adopted the Lloyd George envelope as a state-of-the-art method of data recording. Even so it was difficult for data to be combined for research without huge effort. In 2015 virtually all UK practices use electronic medical records (EMRs) and many are paperless, but that is not the case universally even in industrialised countries.<sup>5</sup> Because general practice is responsible for 80% of healthcare contacts in many countries, data collected during consultations already provides incidence and prevalence information that may be difficult to obtain elsewhere. Where electronic records are available they are also used routinely to manage complex patients, as in this example. On their own, electronic records have been shown to improve the quality of care and flatten inequalities but it is when data can be shared securely and confidentially for research that their value in research rises exponentially.<sup>6</sup> Mackenzie's Institute for Clinical Research was a forerunner of attempts to enable family practices to create a research network and many countries now have research networks in primary care that increase the ways data can be used effectively.<sup>7</sup>

#### RECORD LINKAGE

Record linkage is the process of fusing the data from many events into a story. There are different ways to do this but it can be done safely and securely to improve patient care. Over the past 18 years in Scotland we have been able to take data from people like Mr McGregor and create a register of all diabetics in the country. Practices, hospital clinics, and laboratories share data about patients, making it possible to monitor the quality of care received by every patient,

and to reduce test duplication. Sharing data enables better-quality care at the same or lower cost. For example, the number of lower-limb amputations in Scotland dropped by 40% even as the numbers of diabetic patients more than doubled.<sup>8</sup>

#### ANALYSING DATA

Once data from multiple sources have been combined they may be analysed in more sophisticated ways. Modelling of the Scottish diabetic retinopathy data means that people like Mr McGregor can have the risk of developing blindness assessed and the frequency of testing adjusted according to that risk. A few people at high risk might need review every 3 months and many patients who have an annual check could have the time interval increased to every 2 years.<sup>9</sup> The UK has developed tools to assess risk for many conditions like diabetes, osteoporotic fracture, and cardiovascular disease based on data from the clinical records of millions of patients.<sup>10</sup> In New Zealand that has been taken a step further using a big data approach to update the equation by using what has been recorded about outcomes like stroke and heart attack as it is recorded in patients' records.<sup>11</sup> Big data may be a term with several meanings but we experience its effect every time we shop online and see messages based on our previous choices and those of people like us. For many people this became apparent when an almost unknown senator for Illinois used data in his election campaign for the US presidency. We are already seeing rapid analytical approaches to large volumes of data — big data approaches — which are likely to have a large impact in future.

#### OVERCOMING PROBLEMS

There are technical, governance, and stakeholder issues that need to be addressed to enable optimal use of EMR data however. Most of the technical issues

*"We can all however allow our data to be used and to engage with researchers who can provide that wider vision of more accurate diagnosis, better-targeted treatments, and ultimately better outcomes. Realising that potential will mean that every clinical visit can contribute to advancing medical science."*

around how to extract, process, and use data are normal problems found in many areas of our data-intensive lives. Some countries, like the UK, have strong IT infrastructures designed to enable clinical care, but which are very useful for research. Of course the quality of data varies and its provenance needs to be understood by those who would use it. Major projects like TRANSFoRM are developing tools and services that enable such research to be undertaken across more than one country.<sup>12</sup>

The security and confidentiality issues about using data are the ones that generate most headlines. Most of the examples so far have been for use of de-identified data. Several countries and health systems have evolved ways to achieve this securely while allowing those few people who do not agree that their data be used for research to opt out. An alternative is where people opt in to provide consent to be approached for research. In this case their linked records can help identify those who are eligible studies, as we have done with the SHARE register in Scotland.<sup>13</sup>

There are many stakeholders whose interests need to be considered when research is undertaken using EMR data. The two who need to be effectively engaged in most instances are the data subjects and the data controller. The former is usually the patient but may be another family member or the clinician whose notes create the data. Family physicians act as data controllers for EMR because they legally control and are responsible for the keeping and use of personal information on computer about their patients. Unless the legitimate concerns of patients and their family doctors are adequately addressed then no technical or governance framework can operate effectively, as has been seen in many countries.

### **EMERGENT OPPORTUNITIES**

When the conditions to ensure the safe and confidential use of health data have been met, then a wide range of new research opportunities become possible

and desirable. One exciting development is where patients contribute information and have greater access to their own data. We are just at the beginning of an era when patient-reported outcome measures (PROMS) and other data recorded by patients on their smartphones, pedometers, and monitoring devices become available.<sup>14</sup> One study that combined mobile health applications and a behavioural intervention has already demonstrated a 5 kg weight loss a year after randomisation in overweight Scottish football supporters.<sup>15</sup>

Another major opportunity exists with clinical trials where patients can be randomised on pragmatic criteria to answer the many questions about diagnosis and management in general practice that can only be answered by studies undertaken where most patients seek medical attention. Although there are still problems in making this approach work it is still showing enough promise to pursue as we improve the infrastructure for research.<sup>16</sup>

But perhaps we don't need trials as much as we think we do?<sup>17</sup> For most of the decisions we make under conditions of uncertainty similar options were available to family physicians in practices across the world last year. Their decisions are recorded and the early results are already available. In some cases the same options, decisions, and outcomes have been occurring for 5 or 10 years. If we could agree on how to use that data from patients like Mr McGregor in the UK, Canada, Australia, the Netherlands, and elsewhere to discover which interventions were most helpful for people like him, then we could '*do for medicine what the Atomic Theory has done for chemistry*'.

### **CONCLUSION**

I'd like to finish by repeating a warning given by Sir James to his colleagues in the Institute for Clinical Research at the end of the first year of their work:

*'I must warn you against any immediate expectations of achieving the chief aim*

### **ADDRESS FOR CORRESPONDENCE**

#### **Frank Sullivan**

Department of Family & Community Medicine, University of Toronto, Room 348, 500 University Avenue, Toronto, Ontario M5G 1V7.

**E-mail:** frank.sullivan@nygh.on.ca

*of medicine, the prevention and cure of disease ... But it must be borne in mind that all great enterprises are based on work done by individuals whose past is lost in oblivion ... We must be content that we are playing a necessary part in a great enterprise.'*

Few family physicians want a career in research; they are too busy responding to patients and providing proactive care each day, amounting to 1 million consultations in the UK. We can all however allow our data to be used and to engage with researchers who can provide that wider vision of more accurate diagnosis, better-targeted treatments, and ultimately better outcomes. Realising that potential will mean that every clinical visit can contribute to advancing medical science to fulfil Mackenzie's vision and the College's mission to promote 'science with compassion'.

#### **Frank Sullivan,**

Gordon F Cheesbrough Research Chair and Director of UTOPIAN, Toronto, Canada; FMTU, North York General Hospital, Toronto, Canada; Professor, Department of Family & Community Medicine and Dalla Lana School of Public Health, University of Toronto, Toronto, Canada; Adjunct Scientist Institute for Clinical Evaluative Sciences (ICES), Toronto, Canada; Honorary Professor, University of Dundee, Dundee, UK.

#### **Provenance**

Freely submitted; not externally peer reviewed.

This text is based on the 2015 James Mackenzie Lecture, given in London on 20 November 2015.

#### **©British Journal of General Practice**

This is the full-length article (published online 29 Apr 2016) of an abridged version published in print. Cite this article as: **Br J Gen Pract 2016; DOI: 10.3399/bjgp16X685153**

## REFERENCES

1. Papers and correspondence of Sir James Mackenzie (1853–1925). <http://archiveshub.ac.uk/data/gb237-coll-84> (accessed 4 Apr 2016).
2. Smith BH, Guthrie B, Sullivan FM, Morris AD. Commentary: a thesis that still warrants defence and promotion. *Int J Epidemiol* 2012; **41(6)**: 1518–1522.
3. Macnaughton J. The St Andrews Institute for clinical research: an early experiment in collaboration. *Med Hist* 2002; **46(4)**: 549–568.
4. Wallace P, Delaney B, Sullivan F. Unlocking the research potential of the GP electronic care record. *Br J Gen Pract* 2013; DOI: 10.3399/bjgp13X668023.
5. Schoen C, Osborn R, Squires D, *et al*. A survey of primary care doctors in ten countries shows progress in use of health information technology, less in other areas. *Health Aff (Millwood)* 2012; **31(12)**: 2805–2816.
6. Liyanage H, Correa A, Liaw ST, *et al*. Does informatics enable or inhibit the delivery of patient-centred, coordinated, and quality-assured care: a Delphi study. A contribution of the IMIA Primary Health Care Informatics Working Group. *Yearb Med Inform* 2015; **10(1)**: 22–29.
7. Making primary care better: experiments in front-line medicine. <https://www.youtube.com/watch?v=kLuyXfdDz1A> (accessed 4 Apr 2016).
8. Kennon B, Leese GP, Cochrane L, *et al*. Reduced incidence of lower-extremity amputations in people with diabetes in Scotland: a nationwide study. *Diabetes Care* 2012; **35(12)**: 2588–2590.
9. Looker HC, Nyangoma SO, Cromie DT, *et al*. Predicted impact of extending the screening interval for diabetic retinopathy: the Scottish Diabetic Retinopathy Screening programme. *Diabetologia* 2013; **56(8)**: 1716–1725.
10. Hippisley-Cox J, Coupland C, Robson J, *et al*. Derivation, validation, and evaluation of a new QRISK model to estimate lifetime risk of cardiovascular disease: cohort study using QRisk database. *BMJ* 2010; **341**: c6624.
11. Mehta S, Wells S, Grey C, *et al*. Initiation and maintenance of cardiovascular medications following cardiovascular risk assessment in a large primary care cohort: PREDICT CVD-16. *Eur J Prevent Cardiol* 2014; **21(2)**: 192–202.
12. Ethier JF, Curcin V, Barton A, *et al*. Clinical data integration model. Core interoperability ontology for research using primary care data. *Methods Inf Med* 2015; **54(1)**: 16–23.
13. Sullivan FM, Treweek S, Grant A, *et al*. Improving recruitment to clinical trials with a register of a million patients who agree to the use of their clinical records for research in the Scottish Health Research Register (SHARE). *Trials* 2011; **12(Suppl 1)**: A115.
14. Murphy M, Hollinghurst S, Turner K, Salisbury C. Patient and practitioners' views on the most important outcomes arising from primary care consultations: a qualitative study. *BMC Fam Pract* 2015; **16**: 108.
15. Hunt K, Wyke S, Gray CM, *et al*. A gender-sensitised weight loss and healthy living programme for overweight and obese men delivered by Scottish Premier League football clubs (FFIT): a pragmatic randomised controlled trial. *Lancet* 2014; **383(9924)**: 1211–1221.
16. van Staa TP, Dyson L, McCann G, *et al*. The opportunities and challenges of pragmatic point-of-care randomised trials using routinely collected electronic records: evaluations of two exemplar trials. *Health Technol Assess* 2014; **18(43)**: 1–146.
17. Anglemeyer A, Horvath HT, Bero L. Healthcare outcomes assessed with observational study designs compared with those assessed in randomized trials. *Cochrane Database Syst Rev* 2014; **4**: MR000034.
18. Wilson R. *The beloved physician Sir James Mackenzie*. London: John Murray, 1926.