PERSPECTIVE

Data sharing for discovery

Biomarkers will be essential if research on Parkinson's is to progress, but their discovery depends on scientists sharing data, says **Mark Frasier**.

CHAD BATKA

The current rate of drug development for Parkinson's disease is unprecedented, aided by the discovery of genetic targets and by the creative repurposing of drugs from other diseases. Many of these therapies show promise in being able to slow or halt Parkinson's. Clinical trials of these drugs need reliable tools to assist with patient selection and stratification, and to reveal whether the drug is acting on the underlying disease process. But because of the lack of objective measures of disease progression, it is nearly impossible to determine if a candidate drug is successful.

The absence of validated biomarkers is not due to a lack of information. Rich data sets from observational studies and interventional trials are ripe for biomarker discovery and replication. The problem is that most of the data and biosamples are kept under lock and key. Over the past decade, biomedical researchers have generally become increasingly open to sharing resources, yet this shift in mindset has barely begun for Parkinson's disease. Parkinson's researchers and funders have tended to plan for the short term, focusing on hypothesis-driven studies. The foresight and funding for long-term infrastructure that would enable data sharing have been missing.

Most biomarkers originate from an understanding of disease pathophysiology. Alzheimer's researchers, for example, identified the biomarkers amyloid and tau through knowledge of how these proteins accumulate in diseased brains, transforming Alzheimer's drug development. In Parkinson's disease, imaging markers for dopamine-producing cells have been developed, but these provide only a secondary measure of disease progression and do not reflect an underlying cause. To identify and validate better Parkinson's biomarkers, we must learn more about the disease: the spectrum of symptom presentation and clinical progression, molecular changes before and during the disease, and the effect of these changes on brain structure and function.

Several studies are trying to develop biomarkers for Parkinson's, including the US National Institute of Neurological Disorders and Stroke Parkinson's Disease Biomarkers Program (PDBP)[1], the Michael J. Fox Foundation-sponsored Parkinson's Progression Markers Initiative (PPMI)[2] and the UK Incidence of Cognitive Impairment in Cohorts with Longitudinal Evaluation — Parkinson's Disease (ICICLE-PD)[3]. They collect data from people with the disease and from matched controls; the PDBP and the PPMI make data immediately accessible online for the research community. The information includes physician- and patient-reported clinical assessments, readings from wearable devices, imaging scans, and blood and spinal-fluid samples. Mining these data sets should reveal aspects of Parkinson's and its underlying pathophysiology that could be crucial for biomarker discovery and validation.

Yet many studies, including most clinical trials, archive their data sets on university or company servers, leaving a wealth of knowledge lying dormant. There is little incentive for clinical researchers to make raw data publicly available — the process is strewn with obstacles, including gaining consent from volunteers, resolving ownership of the data and the de-identification of personal information. Universities, companies and their researchers share legitimate concerns about data-sharing efforts: control of data quality and ownership of intellectual property generated on the shared data, and loss of credit if someone else publishes research based on the data. But these factors can be addressed and should not be a barrier to sharing data.

Most people in the United States are willing to share their data anonymously with the research community[4]. Researchers must now buy into that idea. This will entail several specific actions. First, funders must require sharing of the raw data from the project's initiation to ensure that volunteers can give consent and that data infrastructure can be built. Second, researchers must abide by data standards and ontology to help others to use the data. And third, government and database policies must ensure that data are properly de-identified and stored in a safe and secure environment.

THE FORESIGHT AND FUNDING FOR LONG-TERM INFRASTRUCTURE THAT WOULD ENABLE DATA SHARING HAVE BEEN MISSING.

As a prerequisite for these actions, the community must agree on accepted taxonomy and nomenclature. The global non-profit organization the Clinical Data Interchange Standards Consortium is developing terms for data harmonization in medical research, but more effort is needed to standardize the terms that are relevant to Parkinson's[5]. Researchers also need new methods for data integration and analysis to explore pooled data types (such as molecular analyses, imaging data and subjective clinical reports). Scientists should also be able to access standardized data from various studies through a centralized portal. A coordinated public–private partnership that harnesses government financial and intellectual resources could change the research norms and make data sharing the expectation, not the exception.

There are nascent signs that data sharing is gaining traction in research on Parkinson's disease[6]. But, given the stakes, much more should be done. As drug development marches forward, the need for objective, reliable biomarkers grows. These will be discovered most efficiently with infrastructure to support the exploration of existing data. To do this, a collective effort across foundations, governments, universities and industry is needed. As a neutral convener, the Michael J. Fox Foundation invites stakeholders to provide input into this process and help to establish a plan to shift the culture and generate the tools to catalyse biomarker validation. Only by working together in this way will we enable the discovery of new treatments for the millions of people with Parkinson's disease. ■

Mark Frasier *is senior vice-president, research programmes at the Michael J. Fox Foundation for Parkinson's Research in New York.*
e-mail: mfrasier@michaeljfox.org

1. Rosenthal, L. S. *et al. Mov. Disord.* **31,** 915–923 (2016).
2. Parkinson Progression Marker Initiative. *Prog. Neurobiol.* **95,** 629–635 (2011).
3. Yarnall, A. J. *et al. Neurology* **82,** 308–316 (2014).
4. Truven Health Analytics *Health Poll: Data Privacy* (Truven Health Analytics, 2014).
5. Souza, T., Kush, R. & Evans, J. P. *Drug Discov. Today* **12,** 174–181 (2007).
6. Stephenson, D. *et al. J. Parkinsons Dis.* **5,** 581–594 (2015).