Data Sharing: Safely Facilitated Through Shared Responsibility

Not everyone is willing to share, so how can researchers, institutions, and drug manufacturers meet their ethical obligations, without jeopardising their commitments to compliance, copyright, and privacy protection?

As technology continues to advance, patients and healthcare systems alike stand to benefit from the potential of data sharing.

One such benefit is evident within the field of regulatory approval, where real world data are increasingly being used to inform and support drug licensing. Take, for example, the FDA’s approval of an expanded indication for Pfizer’s breast cancer therapy Ibrance (palbociclib), which was based largely on the evidence of shared data from electronic health records and post-marketing reports sourced from real world databases (1). The fact that such information was considered credible enough to help inform a label extension is illustrative of a seismic shift in the industry’s belief in the value of shared data.

Similarly, within the field of clinical trials, peer review journals, sponsors, and funding bodies are increasingly keen to see evidence of accurate statistical analysis. But, despite the growing appetite for data, not everyone is willing to share.
Barriers to Data-Sharing

Institutions involved in clinical trials need to manage sensitive, disparate health data responsibly, while enabling researchers to perform complex analytics using the tools of their choice. Although large volumes of data are being collected, they are not routinely shared within the industry. This can lead to gaps in scientific understanding and may result in valuable insights being missed, which can negatively impact patient outcomes.

It’s therefore an ethical imperative that responsible data sharing is adopted more widely across relevant clinical trials, but without the right policies and products in place, challenges around ownership, rights, responsibility, and regulations can arise at every stage.

Even those institutions that do collect, collate, and store their data effectively, may still stumble at the final hurdle: sharing.

Data can be shared in several ways, including via open or managed access, and policies may have to accommodate periods of exclusivity or embargos. Each permutation will pose its own distinct security and governance issues – both when a trial is active, and once the data have been archived.

Beyond technical barriers, there are also cultural considerations – historically, some investigators have been reluctant to share their research – and practical implications. The merging of multiple datasets can theoretically uncover findings that would never have been spotted in one trial alone, but if there is no common standardisation and systems cannot speak to one another, collaboration becomes impossible.

A Safe Solution

By integrating improved process and advanced technology with an empowered workforce, it is possible to put a compliant and long-term solution to safer data collection and sharing in place.

1. Improved Process

Effective data sharing necessitates that results are credible and accurate, that the confidentiality of research participants is protected, and that interests of individual researchers are preserved. For this to happen, appropriate planning needs to take place, ideally as early as possible during the clinical trial design stage.

Data management models should set out a clear plan for curation, management, and sharing from the outset. These policies need to cover all of the basics – such as gaining patient consent, and ensuring participants are fully aware of how their data will be stored, shared, and used. However, they also need to be sophisticated enough to withstand vetting, be it via an internal ethics committee or external bodies.

Also, data integrity should be preserved (as per 21 CFR Part 11 and the 2018 MHRA Guidance on GXP Data Integrity). ‘Data integrity’ is often associated with ALCOA, which is defined as attributable, legible, contemporaneous, original, and accurate:

- Attributable as being entered by the site personnel and confirmed by the audit trail
- Legible in the site source
- Contemporaneously collected at the time the activity was performed at the site
- Original to the source as confirmed by source data verification (SDV)
- Accurate (i.e., free from errors, complete and within ranges)

The terms ‘attributable’, ‘original’, and ‘accurate’ are particularly connected to electronic data capture (EDC) systems. A robust audit trail, accurately recorded SDV processes, and useful question definition checks enable instant data checks.

Instant data checks highlight any inaccuracies that are set via automatic validation checks when a clinical trial is initially designed in the electronic case report form. These data can then be clarified by the data entry staff straight away. A second step in ‘data cleaning’ is raising queries (another name for a query is a data clarification report [DCR]), followed by the SDV process.

2. Advanced Technology

The right technology should effortlessly facilitate the above processes and allow for every data eventuality – from upload through to search, access, download, and archiving. Researchers need to be able to set access and authorisation levels within distinct data sets, and software should be in place to block unauthorised data copying and prevent data egress.

Of course, the storage and processing of data needs to cater to more than just the specifics of an individual trial. The tech must also allow for the efficient and compliant processing of external data requests and be able to conform to the changing dictates of industry legislation and regulatory changes.
The right technology should also ensure security by adhering to rigorous standards. Any patient data storage solution must adhere to GDPR/HIPAA and should be certified by an external audit to meet the required standard. ISO 27001 can guarantee best practice is being met and that patient data is protected.

3. Empowered Workforce
The necessities of long-term data curation demand the buy-in of a skilled workforce. The collation, curation, categorisation, and secure storage of data and metadata can be a complex undertaking – and more so if research is co-created and/or several stakeholders are involved.

However, a comprehensive data management model enabled by fit-for-purpose, compliant tech can mitigate many of the potential pain points and alleviate administrative pressure, leaving teams free to focus more of their expertise elsewhere. EDCs today are highly sophisticated, while maintaining a focus on user experience, allowing trials to be conducted efficiently and effectively. This makes them feature-rich, offering standardisation, clinical coding, and event-driven email alerts as powerful and flexible tools. Yet, contrary to popular belief, you do not need to be a programmer or undergo extensive training to use an EDC (2). These solutions are designed to be user intuitive, enabling fast data collection to enhance and accelerate research, and include additional features to ensure ease of use:

- Mobile-friendly
- Cloud-based design
- Point-of-care data entry

Most EDC solutions come with customer support services, alleviating potential bottlenecks by taking the support role off study designers.

4. Interacting Systems and APIs
Data flow between multiple systems is often necessary to keep a clinical trial running smoothly. For example, modern EDC systems are capable of interacting with other clinical trial systems, such as randomisation interactive web response system, or clinical trial management system.

To make this data exchange possible, application programming interfaces (APIs) provide the apparatus for communication between these applications. As the EDC system is the central collection area for all data collected during the study and must either push data or allow data to be pulled to all other systems as necessary, the quality of the EDC API is critical.

A well-designed API reduces the workload for researchers and data management staff, replacing manual data handling processes with specifically designed automated processes.

Setting an Industry Standard
EDC solutions offer the security, streamlining and standardisation required to break down data barriers and empower the industry to adopt better policies and processes.

One of these better processes is the implementation of Clinical Data Interchange Standards Consortium (CDISC) standards (3). CDISC is an organisation that creates clarity in clinical research by bringing together a global community of experts to develop and advance data standards of the highest quality by providing data standardisation (e.g., providing set codes for particular data points/questions). CDISC created Clinical Data Acquisition Standards Harmonization (CDASH) coding, which establishes the metadata standards for clinical data collection. Organisations across the industry use CDASH when developing electronic CRFs in EDC systems.

As the Business Case for CDISC Standards states: “Summary notes, standards implemented from the beginning can significantly improve processes in a single clinical study, thus saving time and cost.”

Implementing CDISC standards in EDC systems also makes data planning and data sharing easier on many different levels. Companies using these standards know that their data are comparable and exchangeable. This creates many benefits, such as improved study planning, comparability with other studies, or even giving the opportunity for cross-company collaborations.

Summary
An easy to use EDC system allows users to quickly input, monitor and run reports on subject data in line with internationally recognised ethical and scientific requirements. Coming at a time when there is growing pressure for data transparency from patients, trial sponsors, regulators, and journal editors, EDC systems can unlock the power of sharable data, allowing us to realise the true potential of collaborative research.

References
2. Visit: www.jmir.org/2016/6/e141

Oli Cram has worked providing clinical research solutions since 2006, and in that time he has been a programmer, a tester, test manager, customer support manager, Software-as-a-Service manager, and is now general manager for Veridata EDC. His vision is to use this breadth of knowledge and the support of his experienced team to deliver the next generation of clinical research software solutions. Oli has an Honours degree in Genetics and a Master’s degree in Computer Science.