



Harnessing the untapped potential of legacy data in pharma R&D

By Elif Sikoglu

Clinical trials for a new therapy cost a median of \$41,117 per patient. Costs like this are no surprise to pharma leaders. But during an age of increasing budgetary pressures, drug developers are under pressure to do more with less money and staff. While there are no “simple” answers to this challenge, there is one strategy that offers research and development (R&D) teams a very powerful approach: better leveraging existing legacy data.

Pharmaceutical companies own petabytes of imaging data, generated by in-house research, investigator-initiated studies or clinical trials. This data is valuable and can yield insights that can help researchers better understand disease mechanisms and inform therapeutic approaches. But in many cases, researchers cannot access this important data, as it remains in silos with CROs, investigator labs, or within a specific research group. Even within these silos, inconsistent data management practices hinder data reuse.

The untapped potential of this legacy data, combined with inadequate planning for new data acquisition, can lead to undiscovered insights, redundant work, and additional costs. In recent years, life science innovators have begun to recognize the importance and potential of legacy data. Leading organizations are now investing in strategies and tools that make legacy data easily accessible, standardized, and reusable.

Better insights through data

Historically, life science companies did not utilize the full potential of acquired imaging data due to the associated complexities of working with it. But when analyzed properly, scans from different imaging modalities, such as MRI, CT, etc., can be applied to better characterize disease mechanisms and therapeutic responses, and hence improve predictive models. Such approaches can effectively inform clinical trials as they can support patient-related decisions (i.e., selection and

classification) and novel biomarkers that can become objective study endpoints. With these insights, pharma companies can design trials that are more targeted and efficient, with higher response rates and lower dropout rates.

So, what’s stopping organizations from taking advantage of this potential? The challenge lies in the fact that wrangling imaging data can be particularly difficult, even more so than other types of legacy data. Its size and complexity make it especially likely to remain siloed and untapped after its initial use.

This situation underscores the importance of adopting new approaches to harnessing legacy data. Its value is massive, but accessing it requires rethinking deeply entrenched processes and organizational cultures. The good news is that modern integrated data management tools can automate much of the previously time-consuming work involved in ingesting and standardizing legacy data from multiple sources. When this data is captured and curated in a central repository, researchers can access and utilize it efficiently and collaboratively.

New incoming data approaches

Going forward, research teams should also ensure that newly generated data is optimized for reusability. Pharma companies should establish expectations with CROs and other sources of data acquisition regarding how data should be captured, formatted, and transferred:

- When setting up data acquisition protocols, especially with a CRO, consider the downstream analysis that will be done and ensure the data collection fits the purpose. (e.g., harmonize MR parameters for multisite studies.)
- Ensure acquired data undergoes real-time quality checks and mitigation if needed. Include automated approaches in the process.

- Establish a secure data-sharing technology to enable seamless information flow and collaboration between involved groups and team members.
- Determine the most beneficial data flow frequency between the data collector and the central repository. With the right technology, you can transfer data in batches or in real-time as you capture it. If researchers within the organization find real-time access useful, set rules to automatically send the data as the device captures it.
- In general, proprietary file formats should be avoided at any stage, and researchers should rely on standardized formats. While you can convert some formats, the process is not always straightforward.

Data: The crucial asset

With a vast library of de-siloed, centralized, multimodal legacy data, R&D teams can perform analyses with unprecedented efficiency, leading to more informed decisions about drug development at each stage. The era of uniting and leveraging multimodal data is just getting started, but it promises to change the way that R&D teams work by turning old legacy data combined with newly acquired data into invaluable assets. **DD&D**



Elif Sikoglu

Senior Director of Life Sciences Solutions at Flywheel

Elif Sikoglu is the Senior Director of Life Sciences Solutions at Flywheel, a biomedical research data platform. She holds a Ph.D. in biomedical engineering and was previously the medical director of an imaging CRO. Elif has expertise in using advanced imaging approaches in neuropsychiatric indications and significantly contributed to clinical trials with imaging-supported endpoints across multiple therapeutic areas.



Flywheel

flywheel.io

1015 Glenwood Ave, Suite 300
Minneapolis, MN 55405

(612) 223-7359