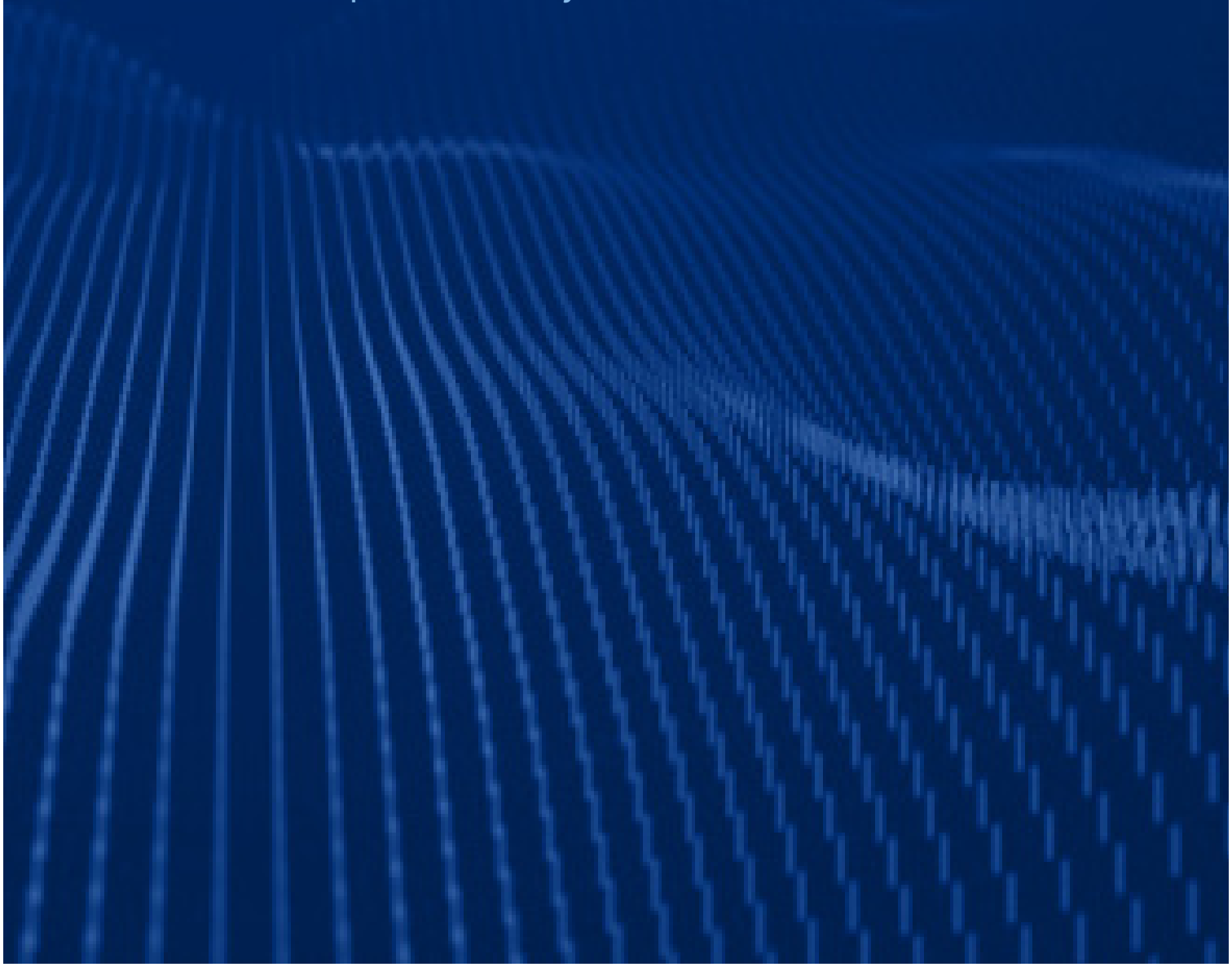
OM1, Inc.
31 St. James Ave Suite #1010
Boston MA, 02116
+1 (888) 324 3899
Info@om1.com
www.om1.com

OM1's Data Automation in Prospective Studies and Registries

Research White Paper | February 2025



Introduction: Addressing the Challenges of Traditional Clinical Research

Clinical research, particularly post-marketing and safety studies, has traditionally been a labor-intensive endeavor, often burdened by **manual effort, high costs, long timelines, and low satisfaction** among sites and patients.

The 21st Century Cures Act, signed into law in 2016, aimed to accelerate medical product development and bring innovations to patients faster. In response, the FDA established a framework for a Real-World Evidence (RWE) Program to evaluate the use of RWE to support new drug indications and post-approval study requirements. This legislative push underscores the growing need for efficient and high-quality real-world data (RWD) for regulatory decision-making. The challenge lies in transforming the traditional, often inefficient, research paradigm to meet these demands.

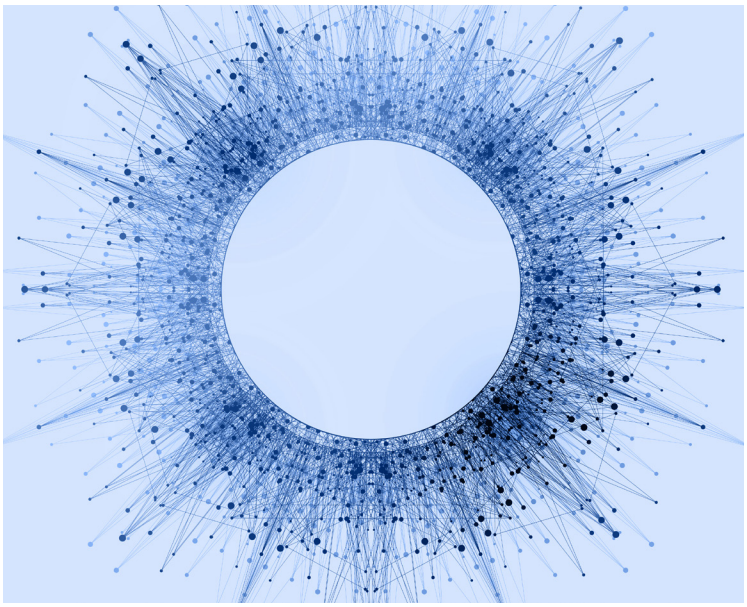
This booklet outlines how OM1's innovative approach, leveraging **data automation and Artificial Intelligence (AI)**, addresses these fundamental problems by streamlining data collection and processing while maintaining traceability and auditability, ensuring regulatory compliance, and enabling more cost-effective, insightful, and patient-centric prospective studies and registries.

OM1's Transformative Approach: Automation and AI

OM1's strategy fundamentally shifts the paradigm of prospective research by employing **automation to make studies more predictable and reliable, reduce marginal costs, and achieve massive cumulative savings**. This is achieved by combining **passive data collection** (using existing systems of record) with **active data collection** (for data not routinely captured).

Passive Data Collection

OM1 emphasizes the direct, consented retrieval of Electronic Health Record (EHR) data. EHRs are a primary source of RWD, containing a patient's medical history, diagnoses, treatment plans, immunization dates, allergies, radiology images, pharmacy records, and laboratory results. OM1 connects to central EHRs and ancillary systems of record, such as Laboratory Information Management Systems (LIMS), Radiology Information Management Systems (RIMS), and Pathology systems. Additionally, **linked data from various sources** like medical and pharmacy claims, social determinants of health (SDoH), and mortality data may also be integrated by leveraging common token systems. This extensive data linkage increases the breadth and depth of information on individual patients over time.



Active Data Collection

When essential data elements are not routinely collected in clinical records, **ancillary data collection** is required. OM1's approach limits active data collection to the minimum necessary, making studies and registries simpler for providers and patients. This includes collecting:

- Electronic Clinical Outcome Assessments (eCOAs).
- Electronic Patient-Reported Outcomes (ePROs).
- Specialized clinical and laboratory assessments.
- Virtual follow-up assessments through a virtual or central site maintained by OM1 where applicable.

In other words, **the most efficient and least burdensome model for data collection for prospective studies and registries leverages: passive data collection of existing data sources using automated data collection methods (directly interfacing with health records) and AI-based abstraction of unstructured data and integrated active data collection for those data elements not routinely collected** under standard of care such as specific eCOAs, ePROs or specialized tests. In uncommon or rare diseases or long-term follow-up studies, direct patient enrollment through a virtual center where a patient provides consent for OM1 to retrieve their data from any healthcare facility that they may visit further improves this model.

Leveraging AI for Data Automation

AI plays a crucial role in enhancing data completeness and study efficiency:

- **OM1 uses AI to map and process structured data** in multiple formats from any number of clinical centers into a common data file. These transformations are critical to having a common data model with common terminology. Each transformation is fully logged to maintain traceability. AI also aids in quality control, cohort generation and analytics.
- **OM1 extensively leverages AI-enabled and clinician validated text extraction** to extract data from unstructured sources like physician notes, images, and radiology/pathology reports, converting them into structured data. This automates collection of the significant amount of key clinical data often residing in unstructured formats within EHRs that would typically require manual abstraction by a clinical site.
- **Digital Phenotyping (PhenOM®)**, powered by AI, facilitates significant new capabilities from identification (for recruitment) of patient cohorts within EMR data (e.g. rare diseases) to predicting baseline adverse event frequencies.

Meeting Regulatory Standards with Automated Studies and Registries

The FDA's RWE Program requires that RWD be **"fit-for-use"** in regulatory decision-making, emphasizing **relevance and reliability**. Relevance pertains to the availability of data for key study variables (exposures, outcomes, covariates) and sufficient numbers of representative patients. Reliability encompasses accuracy, completeness, provenance, and traceability. OM1's automated study and registry programs are designed to meet these stringent requirements:

- **Data Quality and Standardization**: OM1-managed studies collect structured and predefined data elements, offering longitudinal, curated data about defined patient populations. The use of **common data elements** promotes standardized, consistent, and universal data collection, facilitating comparison and linkage with other sources. This approach ensures conformance with FDA's requirements for submitting study data in applicable drug submissions.

- **Curation and Transformation**: Data curation applies standards to source data, such as coding for adverse events or disease progression. Automated processes for data transformation, extraction, cleansing, and integration into a Common Data Model (CDM) are meticulously documented, including justifications for approaches used to reconcile challenges like inconsistent coding, changes in practices (e.g., ICD-9 to ICD-10 codes), or missing information. This documentation is crucial for FDA review.

- **Traceability and Auditability**: OM1's systems maintain traceability of data from analysis results back to source data and audit records, often with electronic signatures, which is vital for verification during FDA inspections.

- **Prospective Planning and Consultation**: Sponsors are encouraged to **consult with the appropriate FDA review division early** in the process and submit protocols and statistical analysis plans **before conducting studies** that include registry data. This pre-planning ensures that the registry's design, including target population definition, data elements, linkage strategies, and outcome validation methods, aligns with regulatory expectations.

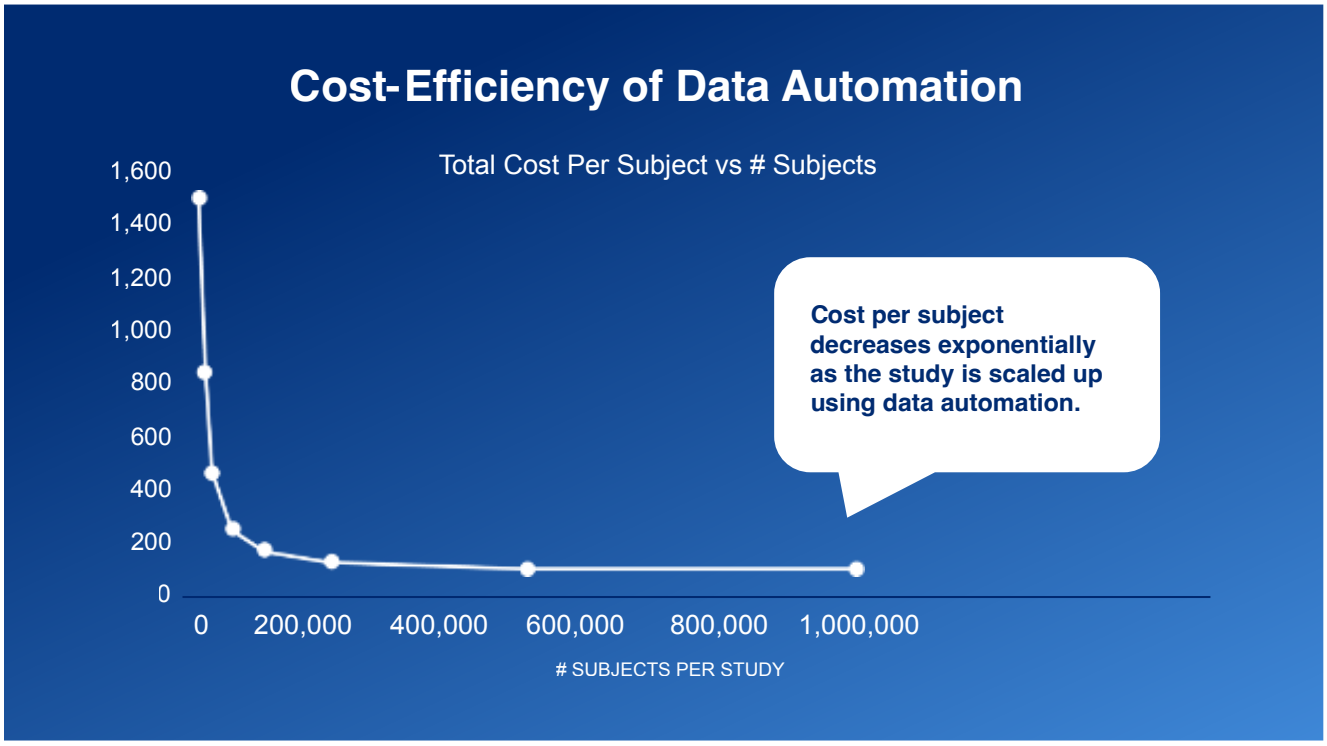
Cost Efficiency and Diverse Purposes of Registries

Cost Reduction Curve Based on OM1 Studies

One of the most significant advantages of data automation is its impact on cost. As shown in OM1's studies, the **cost per subject decreases exponentially as the study scales up** using data automation. This efficiency is a direct result of automating processes that traditionally require manual effort, such as data extraction, cleaning, and integration. For example, studies with hundreds of thousands to a million subjects become far more cost-effective with automation compared to traditional methods.

Figure 1 below shows a curve comparing total number of subjects in OM1 studies versus the average cost per subject. Study sizes range from 1500 patients to more than one million. As shown, as the number of patients or subjects increases, the unit costs per subject reduce dramatically. At numbers below 1000 patients, costs are closer to traditional methods but site satisfaction is higher as manual effort is still a major component of investigator site burden. In these smaller studies, OM1 achieves further cost efficiencies by leveraging its virtual center approach to obtaining medical records as well as supporting active data collection where feasible.

Figure 1

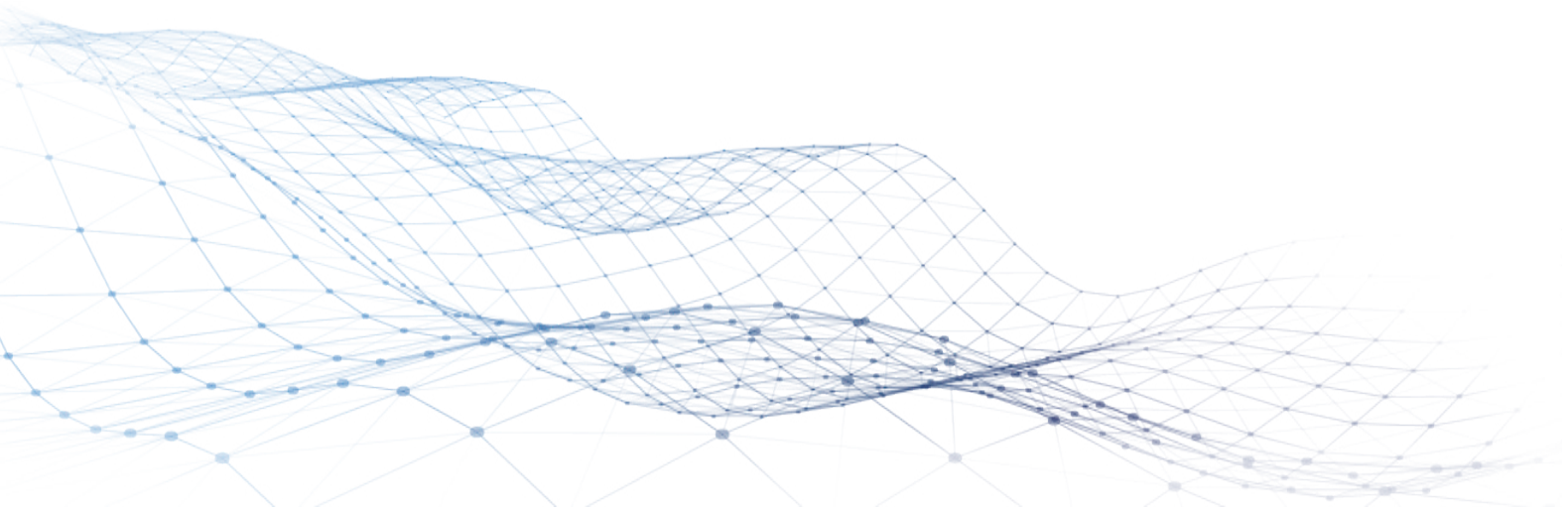


The second advantage of reducing manual effort is a significant increase in site satisfaction with their participation in automated studies and registries. OM1 measures Net Promoter Scores (NPS) in its studies from site research coordinators. As shown in some of the case examples below, OM1 automated registries and studies have **NPS typically above 70** which indicates an extremely high satisfaction level.

Diverse Purposes of Prospective Studies and Registries

Programs managed by OM1 leveraging automation and AI serve multiple critical purposes throughout the medical product lifecycle, from early development to post-market surveillance. OM1's automated approach enhances the utility for all these applications:

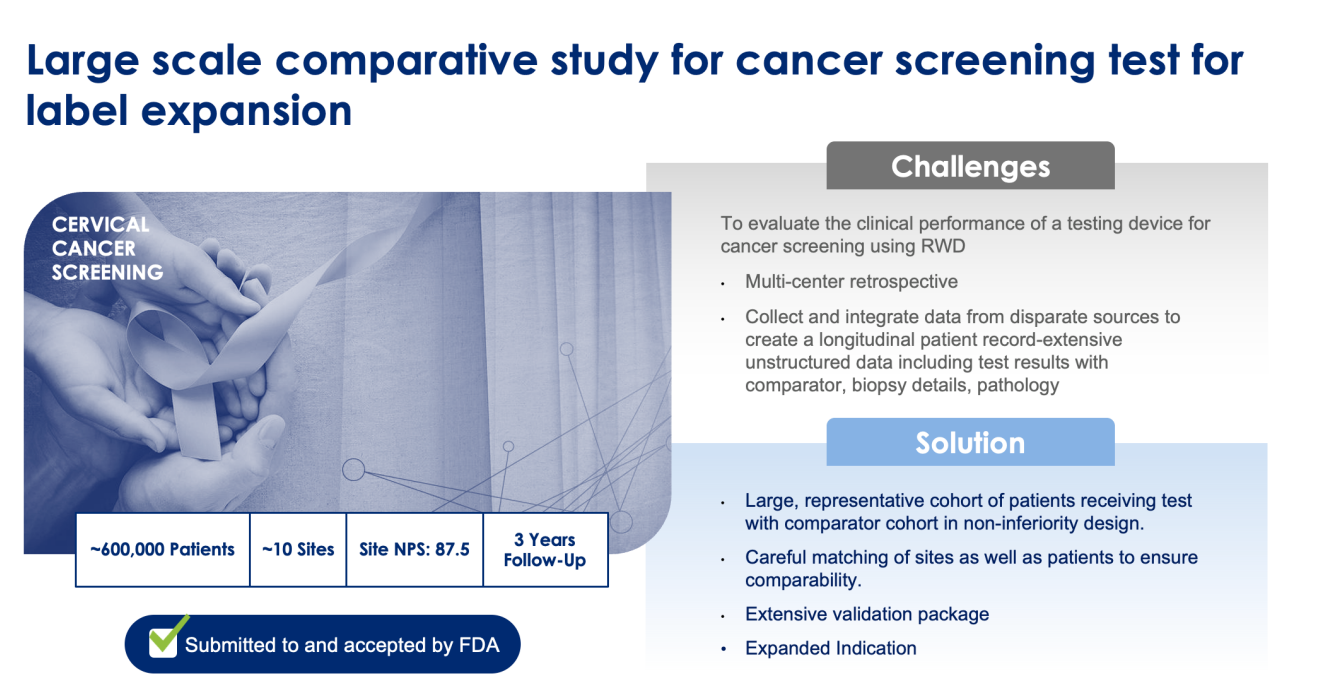
- **Safety**: Post-approval safety studies and registries are invaluable for measuring or monitoring safety and harm associated with specific products and treatments, including conducting comparative evaluations of safety. They can systematically collect data on adverse events and their incidence, addressing limitations of spontaneous reporting systems. For instance, a registry can evaluate safety signals identified from other sources or assess factors affecting risk like dose and timing.
- **Effectiveness**: Post-market studies and registries help determine clinical effectiveness or cost-effectiveness in real-world clinical practice. They are well-suited to evaluate drugs received during routine medical practice and provide information on long-term efficacy outcomes. They can address gaps in generalizability from clinical trials by including more diverse and representative patient populations (e.g., older patients, different ethnicities, those with comorbidities) who might be underrepresented in traditional trials.
- **Natural History Studies**: They are common platforms for natural history studies aimed at understanding disease progression, identifying demographic, genetic, environmental, and treatment variables that correlate with disease development and outcomes.
- **Supporting Clinical Trials and Regulatory Needs**: OM1 provides these capabilities for external controls for interventional trials, to support or satisfy post-approval study requirements, or to provide a framework for trials embedded within registries. They also enable a cost-effective approach to long-term follow-up studies, including for small populations (e.g. rare diseases) with long durations (e.g. gene and cell therapies). Some sponsors are beginning to explore using OM1's systems for automated collection of passive data as a routine part of randomized clinical trials to reduce manual entry and costs.



OM1 Key Use Cases and Accomplishments

OM1's data automation and AI capabilities support a wide range of use cases, demonstrating significant advantages over traditional methods. The following are examples from OM1 experience over the last two years:

1) A very large, multi-center real world evidence study for a new label indication in oncology diagnostics: OM1 has processed structured and unstructured data from 10 major health systems with over 600,000 cases now submitted to the Food and Drug Administration for a new label indication. The processing of these cases leverages 37 AI algorithms for processing unstructured data endpoints that were validated as part of the submission including clinical history, biopsy results and so forth. The data underwent 'fit for purpose' review by FDA as the first step in submission and determined to be 'relevant and reliable'.



2) A large, multi-center prospective comparative effectiveness and safety study evaluating a neurological disease in 10,000 patients eligible for infusion therapy: OM1 is collecting EMR data (including imaging reports and clinical notes) and patient-reported outcomes data and neurological testing (using centralized raters) with five years follow-up. Patients are enrolled or referred by sites and over 50% are followed through an OM1 virtual center.

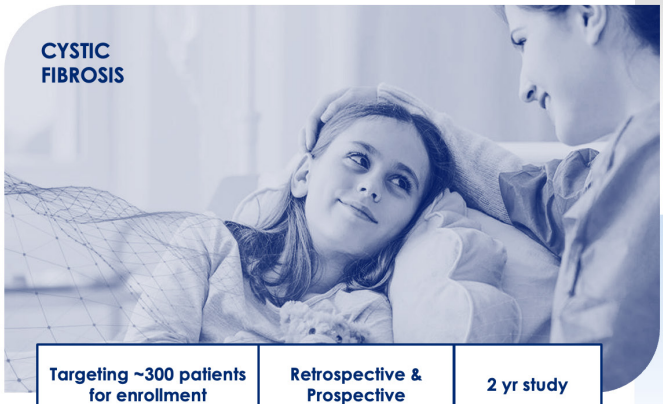
3) A multi-center, prospective study leveraging 30+ centers of excellence and 'referral-only sites' and an OM1 central site responsible for clinical outcomes assessments, PRO, and EHR data collection over 3 years: This is in a rare and potentially fatal dermatologic condition where patients need to be diagnostically verified and enrolled prior to the first initiation of an infused biologic.

4) A multi-center, retrospective and prospective registry in multiple sclerosis where data is collected using automation from EMRs, clinicians and patients (including diary and ePROs). Both clinicians and patients access their own dashboard with summary and benchmarking information.

5) A multi-center retrospective comparative effectiveness study of a pharmaceutical product used in post-transplant infectious disease in 3,000 patients. OM1 is collecting data, including EMR and specialized third party laboratory data, from major transplant centers in the U.S. This study fully leverages automated data collection and processing of unstructured clinical management, test reports and outcomes data.

6) A multi-center retrospective and prospective study identifying and following 300 patients with a rare pulmonary disease, including obtaining genetic data and clinical narratives for case adjudication.

Describing burden of disease in cystic fibrosis



Targeting ~300 patients for enrollment	Retrospective & Prospective	2 yr study

Challenges

- CF is associated with significantly lower quality of life and a high disease burden
- Evaluate the change in patient-reported and caregiver-reported burden of disease after initiating CF treatment.

Solution

- Conducted an observational study to evaluate the real-world effectiveness of treatment on patient and caregivers' burden of illness
- PROs and Caregiver Reporter Outcomes were collected every 2 months
- Described select clinical outcomes

7) A large, multi-center FDA-mandated post-marketing commitment (PMC) study following 50,000 patients tested for colorectal disease. Data from EMRs and laboratories collected by OM1 is submitted quarterly to the FDA. Site satisfaction is very high.

Support post market commitments and expand label for a lower age group

~50,000 total pts	~10 sites	5 years	Site NPS 71

Challenges

Accessing, processing and harmonizing complex and extensive RWD to support post-market commitments and label expansion needs is costly for sponsors and burdensome for providers. RWD is needed to evaluate the effectiveness of a screening device for colorectal cancer (CRC) in high-risk patients.

Method

We are conducting an observational, prospective multi-center study to collect data, including:

- EMR, laboratory, pathology and radiology, and in vitro diagnostic data
- Structured and unstructured data (e.g., pathology reports)
- From Integrated Delivery Networks (IDNs), hospital systems and large medical practices
- To submit for regulatory approval for an expanded indication

8) A multi-center breast cancer screening registry leveraging more than 60 facilities and one million patients with retrospective and prospective data collection. Data is collected from EMR, radiology image management systems, and tumor registries with 8 years of follow-up. The study has focused on effectiveness of specific screening patterns and publications have directly impacted guidelines.

9) A pilot study using digital phenotyping to identify patients with a rare disease for study inclusion. The complete EHR records for an entire health system are screened using the PhenOM® API and identified patients with a high risk of a specific lipid storage disease are recruited for enrollment into the study.

Conclusion

The landscape of clinical research is undergoing a fundamental transformation, driven by the imperative for more efficient and comprehensive evidence generation, particularly from real-world sources. The traditional, manual approaches are no longer sustainable for meeting the demands of modern medicine and regulatory requirements.

OM1's approach, integrating data automation and AI, offers a robust solution by dramatically reducing manual effort and costs, enabling scalable and high-quality data collection. By seamlessly combining passive data from existing systems with active, patient-centric data collection like PROs and eCOAs, these automated studies and registries generate rich, longitudinal datasets. This comprehensive data capture, coupled with stringent quality control and documentation practices, ensures that the data meet the relevance and reliability standards required for FDA regulatory submissions. The demonstrated exponential reduction in cost per subject as studies scale up further underscores the economic viability of this new paradigm.

Ultimately, this revolution in prospective studies and registries means a future where every clinical trial can have a low-cost, real-world, long-term follow-up study, leading to improved time to data and enhanced patient outcomes. The continued evolution of data automation and AI will further solidify their role as indispensable tools in advancing medical product development and patient care.