Leveraging Data Science & Analytics to Optimize Clinical Trial Execution

Melanie Hullings

Director Clinical Data Analytics & Programming

Formation Bio

Melanie Hullings

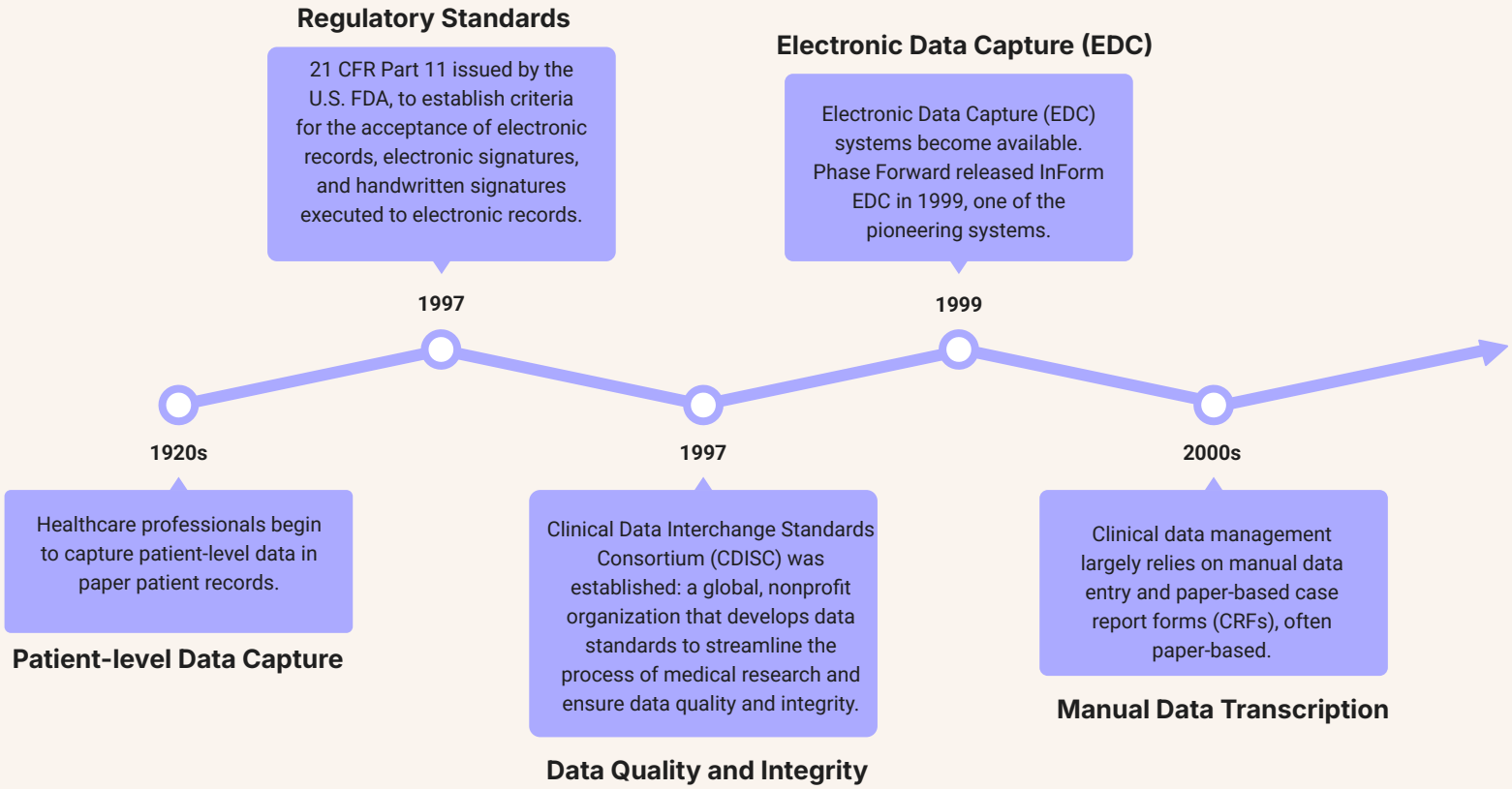
Director, Clinical Data
Analytics & Programming

Agenda

- Historical Context & Transformation
- Overcoming Technological Challenges
- Impact & Results
- Future Directions

Historical Context & Transformation

History of Clinical Data Management (CDM)



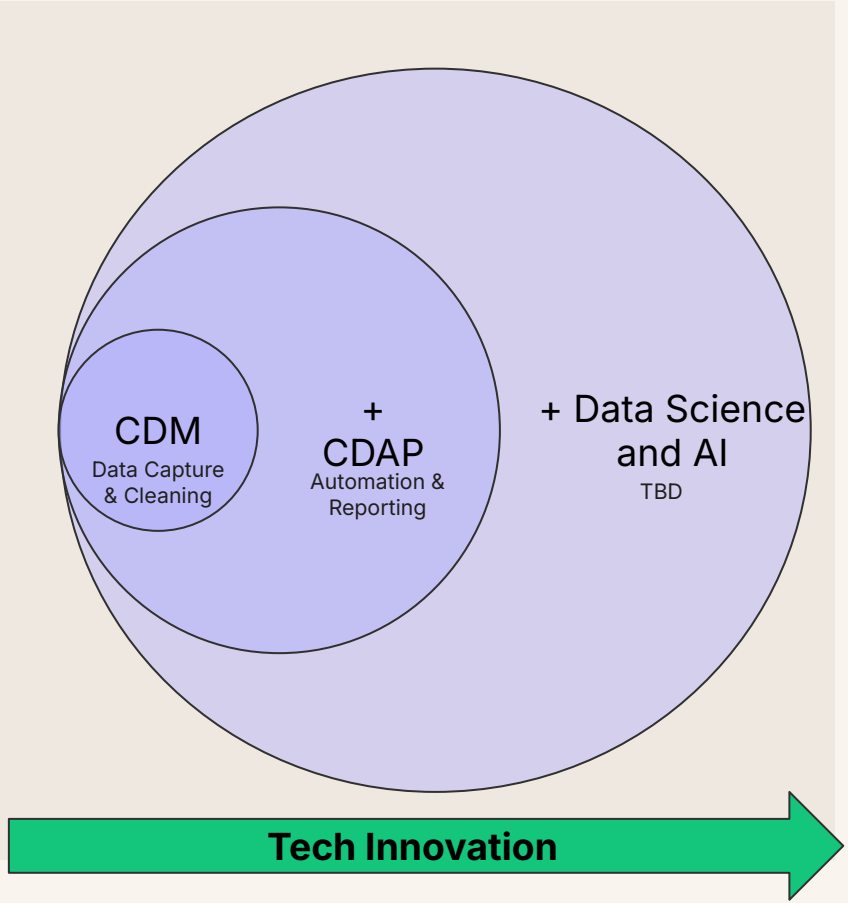
Reference: https://novotech-cro.com/faq/clinical-data-management-past-present-and-future

Proliferation of data capture systems introduces opportunity and challenges

Drivers	Challenges	Solutions
• Digital transformation • Cloud computing • Big data analytics • Patient-generated data • Compliance and regulation	• Data management • Storage • Security • Privacy • Quality	• Data governance • Advanced analytics • Scalable infrastructure • Data minimization • Security measures • Training and awareness

Long-term vision to adopting advanced analytics

- Started with a vision to automate clinical trial execution and data management
- Over the last 3 years, we have built out operational & technical infrastructure to support automation of clinical trial data cleaning and reporting
- Clinical Data Analytics and Programming (CDAP) team evolved from Clinical Data Management (CDM)
- We are now expanding our infrastructure and capabilities for data science to support end-to-end drug development



A new programming function was created at Formation Bio from CDM

Typical Biometrics Team

- Biostatistics
- Clinical Data Management
- Statistical Programming

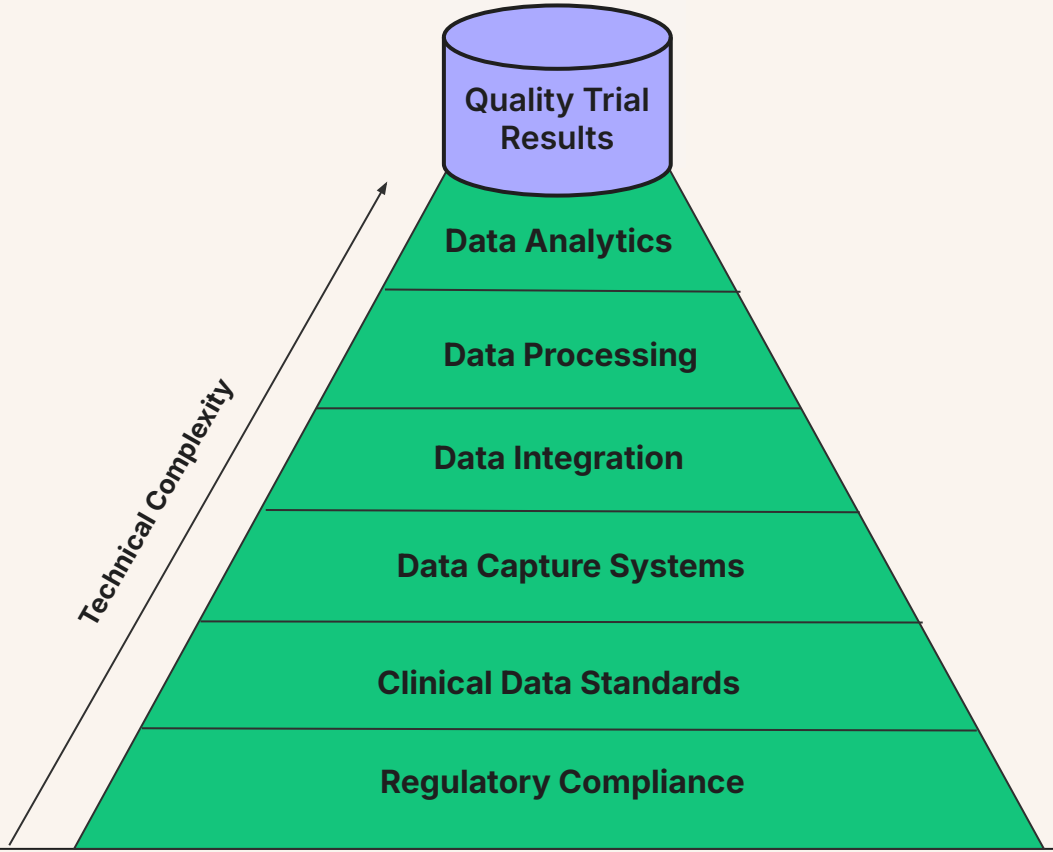
New Function

Clinical Data Analytics & Programming (CDAP)

- Advanced Analytics
- Data Science

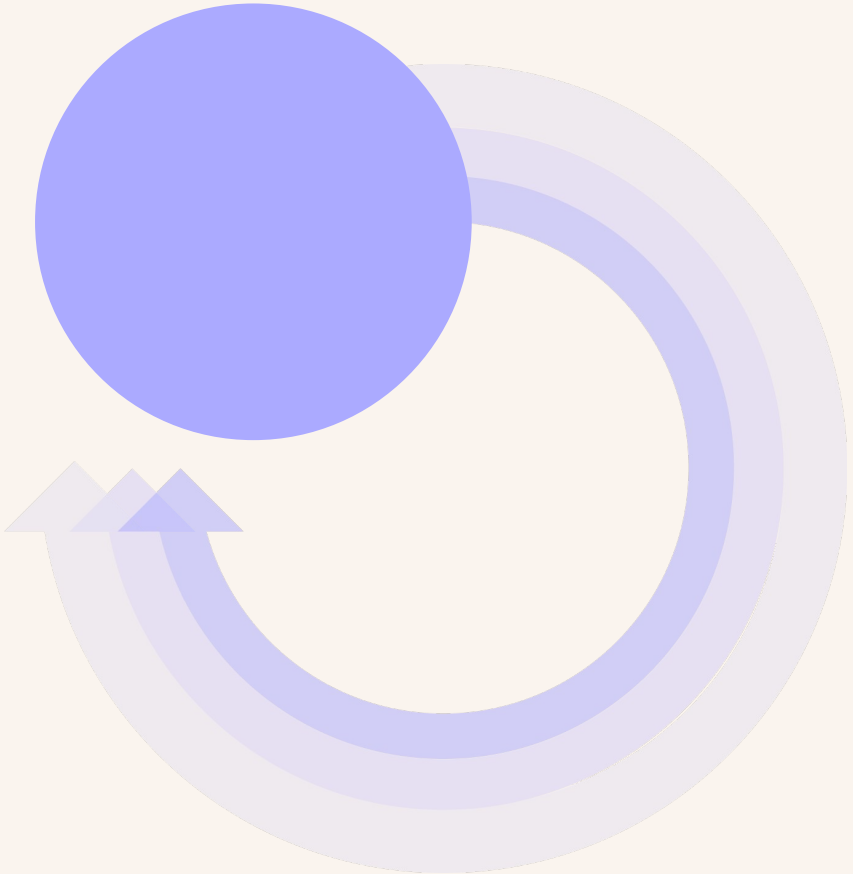
Overcoming Technological Challenges

CDMs build the data foundation for Clinical Data Analysts (CDAs)

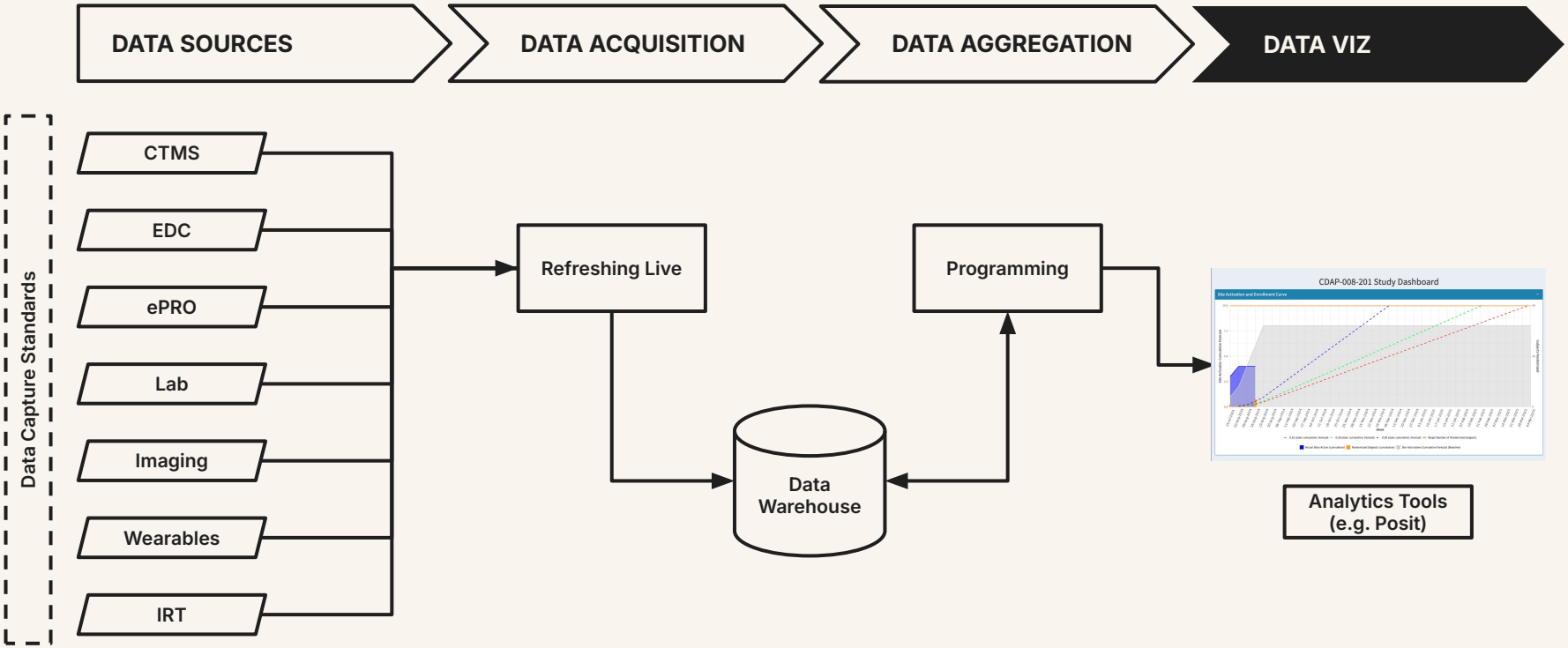


CDM<>CDA Collaborative Model

- **People and process**
  - Maintain focus on data quality and integrity
  - Agile, iterative team operations
- **Innovative data capture**
  - Data standards
  - Site and patient-centric data capture systems
- **Adopt advanced technologies**
  - Open-source programming
  - Machine/deep learning
  - AI



Building a standardized, validated clinical data warehouse was crucial to unlock automated reporting

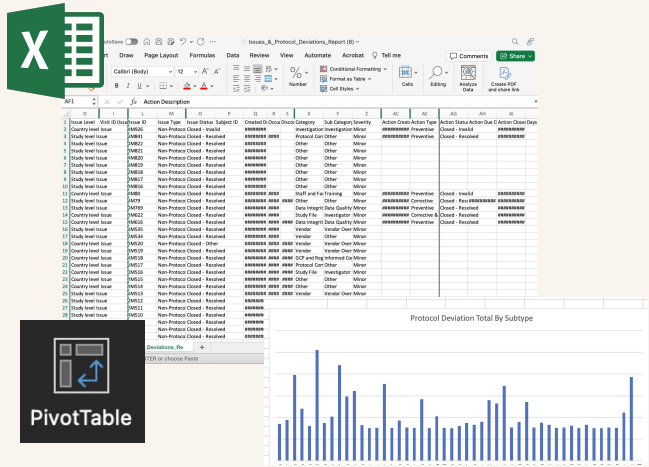


Impact & Results

Initially utilized Shinyapps.io to demonstrate utility of R programming for automating CDM workflows

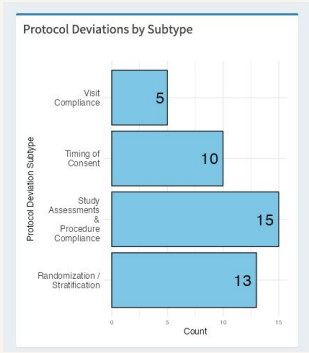
Export data

5+ Steps



VS

1 Click

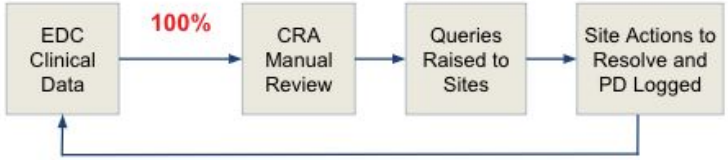


Piloted R program to detect protocol deviations

Process Improvement

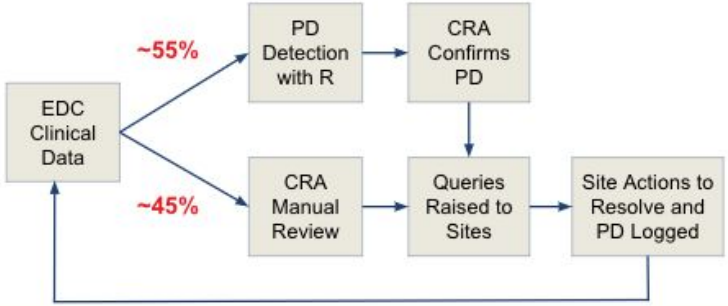
Industry Standard:

- relies on manual review by CRAs to identify deviations from clinical trial protocol requirements



Program Assisted:

- more efficient and ensures holistic issue identification



Results

- Script identified 33 PDs that were not identified by CRAs
- 63% of PDs identified by script were false positives
- False positives have value in identifying data abnormalities

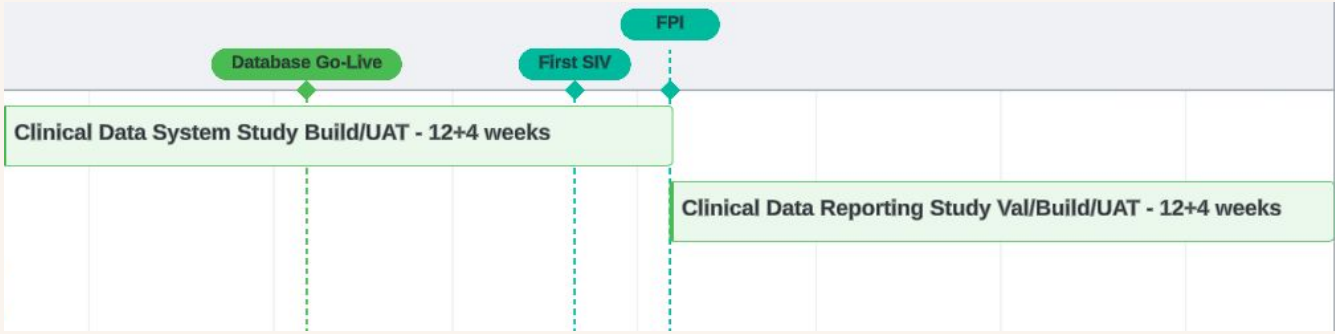
Conclusion

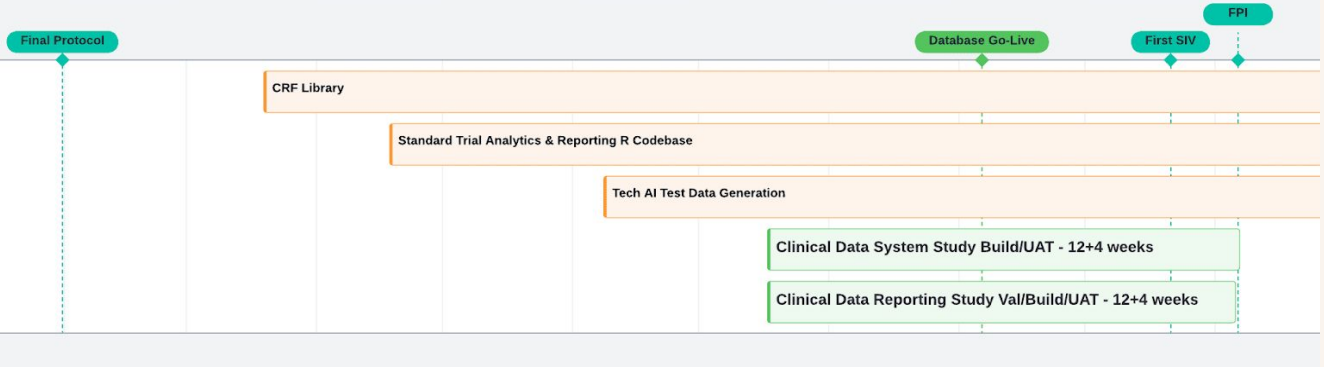
- R scripts can be used to programmatically detect issues in clinical trial datasets including PDs and participant eligibility
- Programmatic issue detection has the potential to decrease manual review burden of study teams, as well as cost, and improve issue identification and clinical trial data quality overall
- Future directions include refining scripts to make issue identification more accurate and expanding the number of data issues that scripts are detecting

CRA = Clinical Research Assistant
EDC = Electronic Data Capture
PD = Protocol Deviation

Utilizing R has allowed us to shift timelines significantly left on reporting

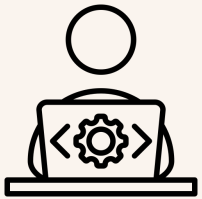
In the past, reporting work could not begin until the database build was complete which led to just-in-time reporting and lack of real-time insight into the study.





We have built out standard data capture, reporting, and tech infrastructure to shift study reporting left during study startup so that reports are available by FPI.

Results of embedding programming on the clinical study team



Automate away 100+ hours of manual data cleaning and review to identify potential data issues, including deviations

Improve speed and accuracy of clinical insights by at least 50% so the study team can focus on resolving issues rather than finding them

Improve collaboration with the entire clinical study team having access to real-time reports

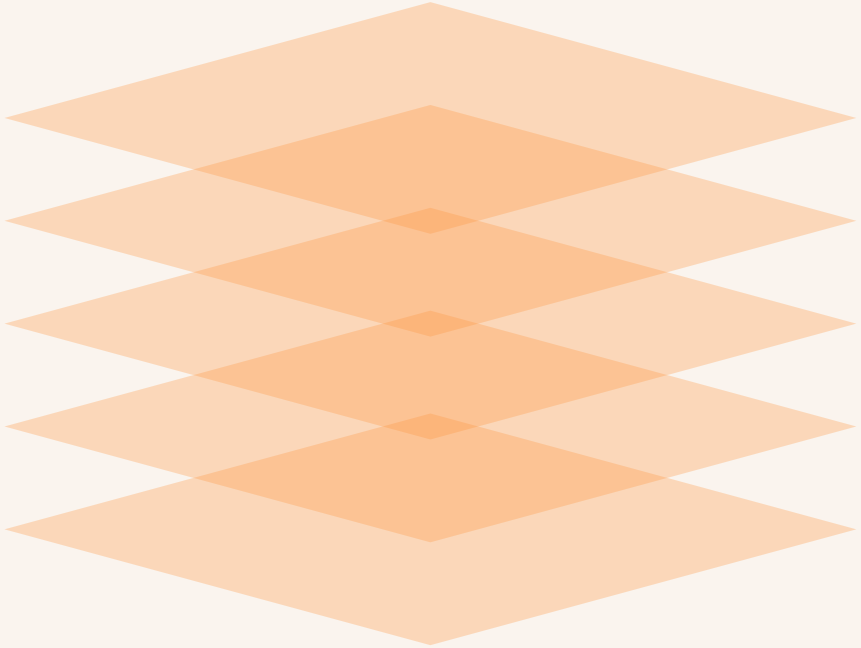
Path Forward

Upskilling company-wide with new technology

- Open-source programming continuing education
  - Applying industry collaborations & conference learnings
  - R/Posit training & workshops
  - Internal R Users Working Group
- Expanding AI toolkit to support programming
  - Custom ChatGPT workshops
  - AI-assisted workflows (e.g. test data generation)
  - LLM tools (e.g. Copilot, chattr)

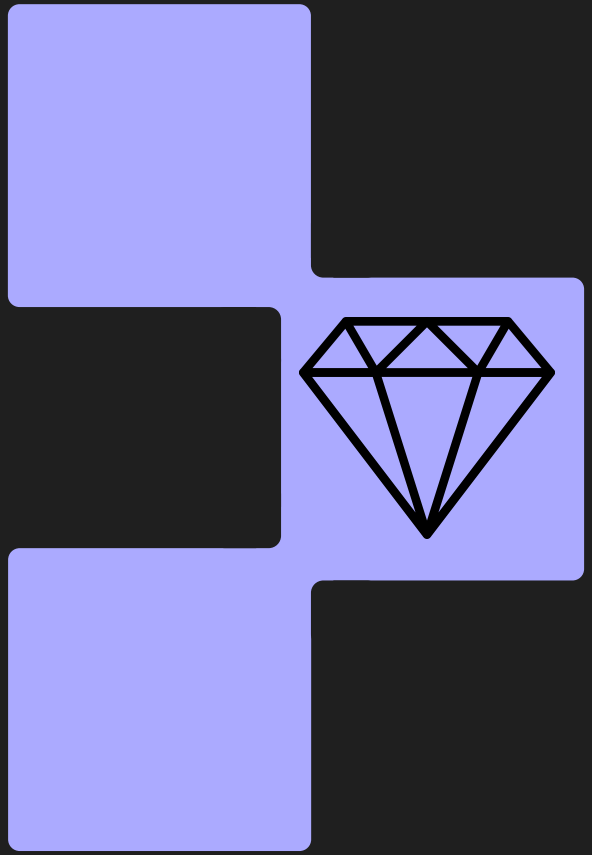
Opportunities for AI in Biometrics

- Areas of opportunity:
  - Documentation & TMF
  - Validation & UAT
  - CRF and edit check configuration
  - Data analysis & visualization
  - New vendor capabilities
  - Statistical programming
  - and so many more...



What AI means for the Biometrics teams of the future

- All roles supported by ML and AI to become more efficient
- Opportunities to streamline and automate processes to become more efficient
- Ability to triage issues immediately to significantly improve data quality
- Proactively informing study operations and success
- Continuous learning



Questions?

Thank You!

Feel free to contact us:

Melanie Hullings (melanie@formation.bio)