

Paper SI05

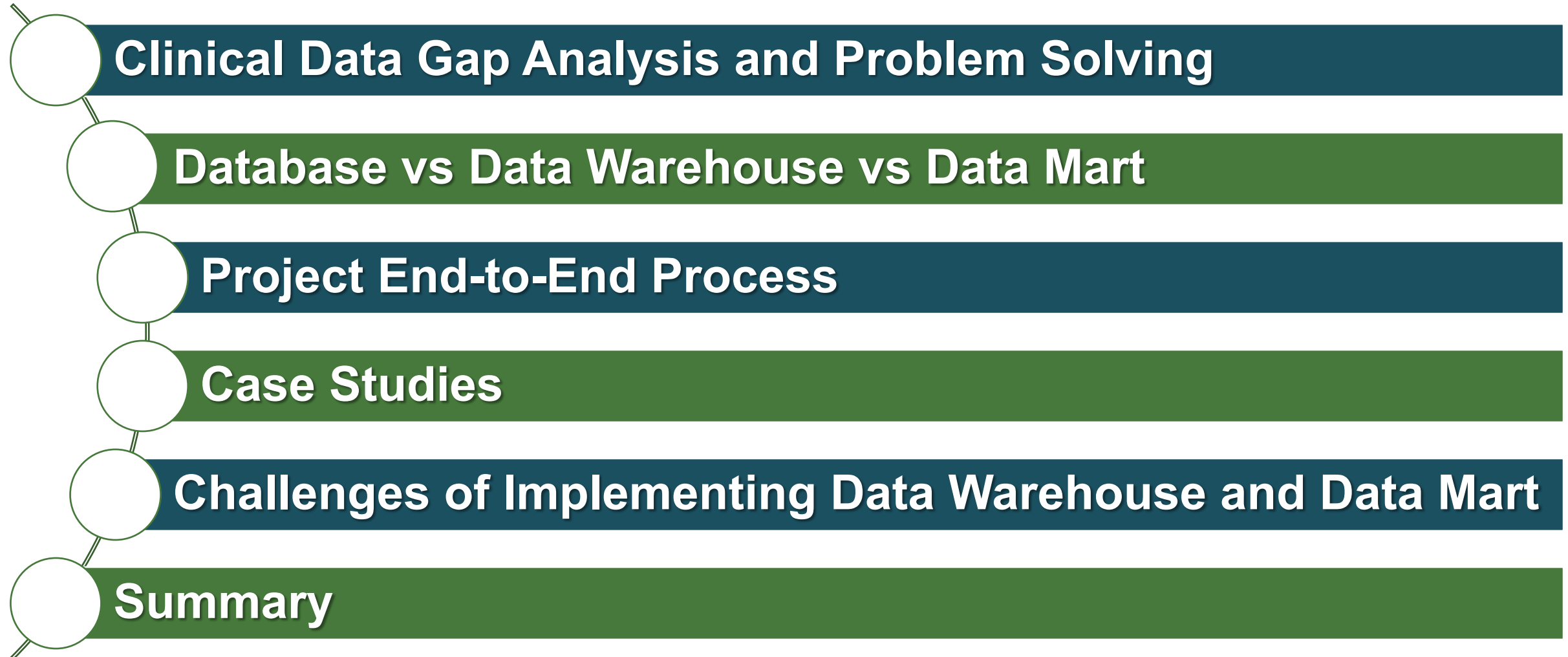
Clinical Data Standardization and Integration Using Data Warehouse and Data Mart Process

Clio Wu

Director, Statistical Programming, Chinook Therapeutics, Inc.

May 3, 2022

Overview



Gap Analysis



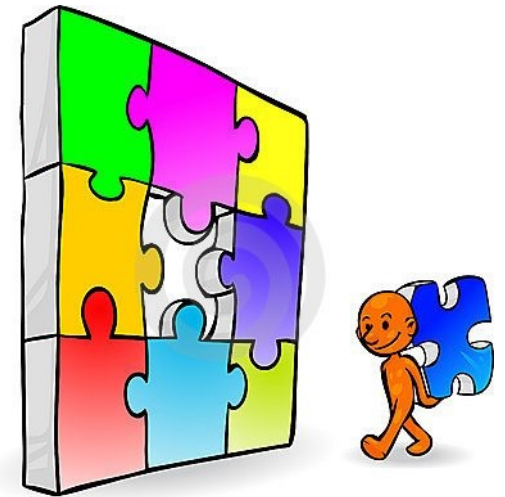
Problem Solving



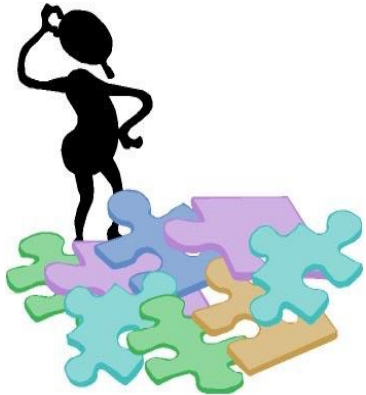
- Explosive growth of pipelines & outsourced studies
- Which is the most current use of data across studies

Data Warehouse

- ✓ Data from different studies are organized, stored and made available
- ✓ Always have the most current version of data with the most current version of programs/utilities



Problem Solving



- Integrated studies, pool data, re-standardize data repeatedly
- Code-list and dictionaries version issues
- Programming R&Rs

Data Mart

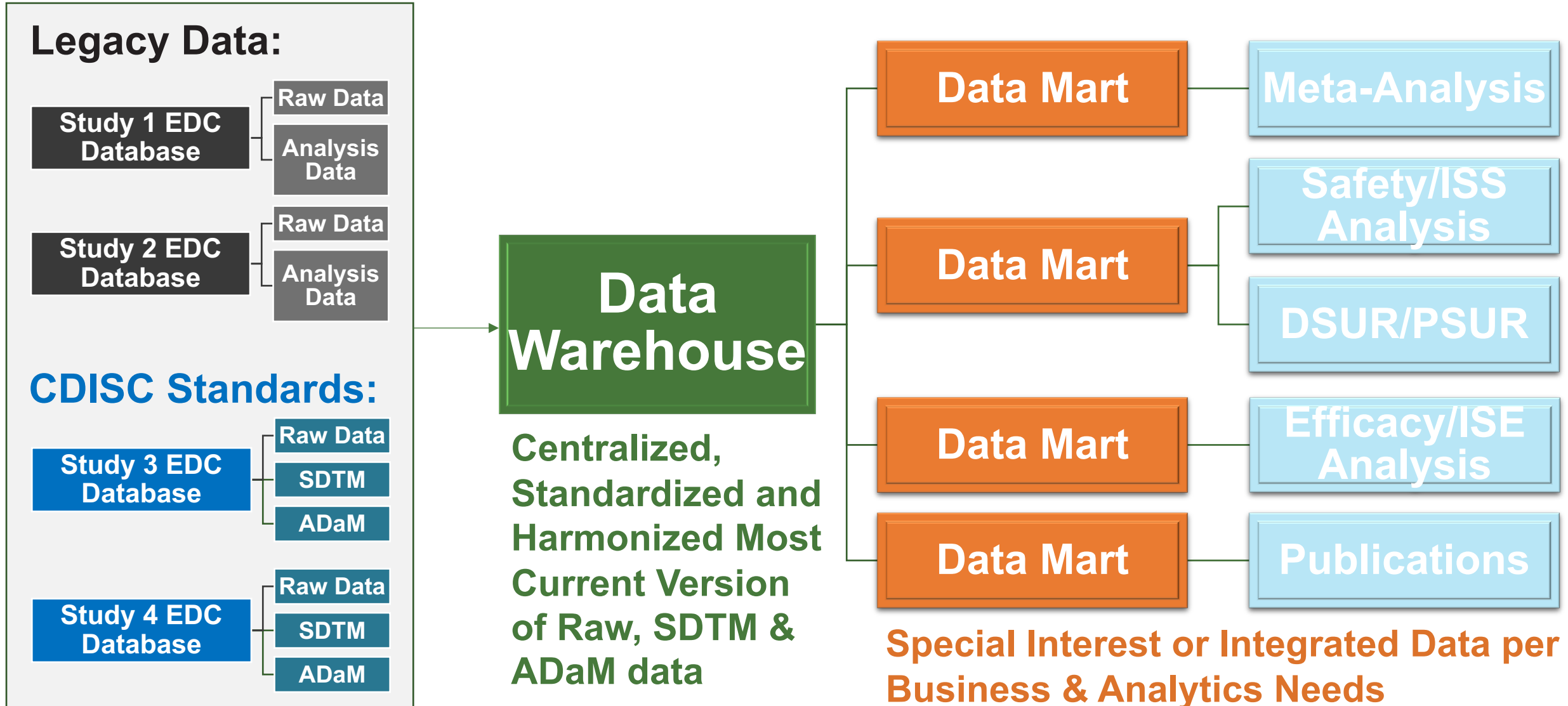
- ✓ Special collections of data organized for a special purpose
- ✓ Transformation of the data into a single standardized data model to facilitate integrated data analyses
- ✓ Get data in a standard way, the integration programmers are not necessary to be the study lead programmers



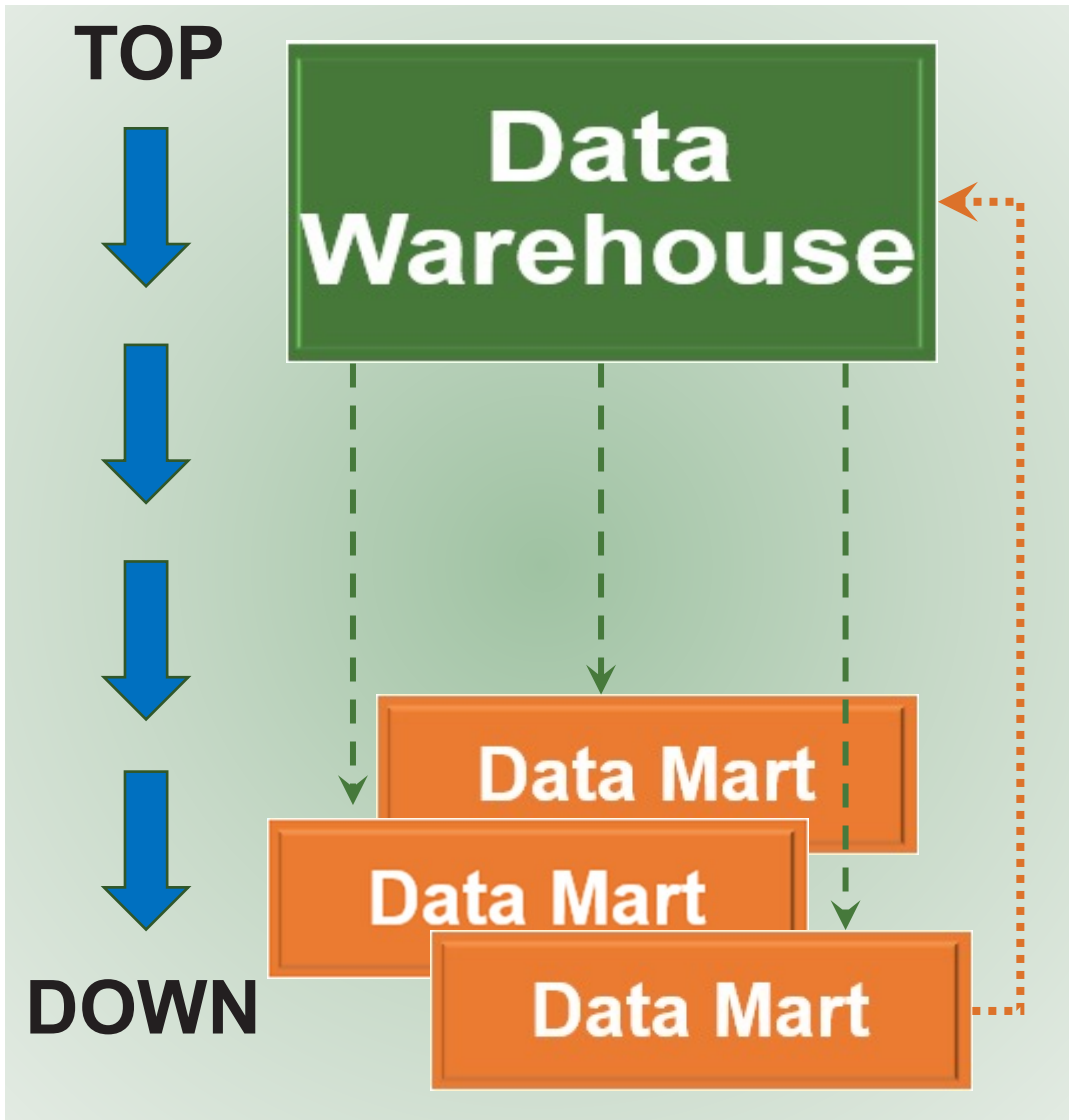
Objectives

- ❑ To develop a system that collects, organizes, pools, and maintains the most current version of raw, SDTM, and ADaM datasets across studies.
- ❑ To facilitate the rapid production of statistical output in an efficient and accurate manner.
- ❑ The “**data warehouse**” aspect of this process will collect, covert and organize all current version of raw, legacy and/or CDISC SDTM and ADaM datasets automatically.
- ❑ The “**data mart**” aspect of this process will collect and/or create pooled data across multiple studies.

Database vs. Data Warehouse vs. Data Mart



Data Warehouse vs. Data Mart Approach



■ Top-Down Approach:

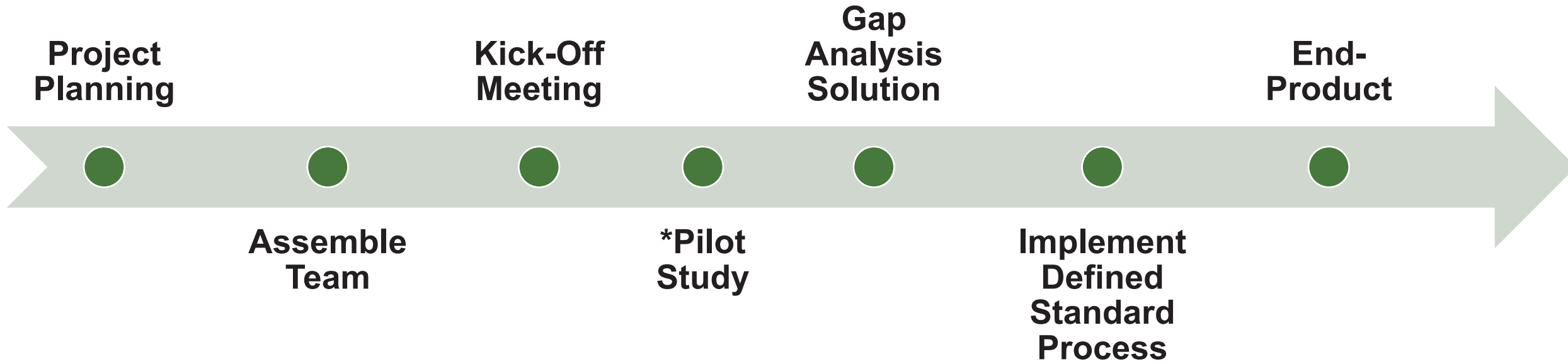
- No timeline, cost and resources concerns
- Full scale of study & data standardization and harmonization



■ Special Interest & Analysis Driven Approach:

- Timeline, cost and resources concerns
- Start with selected study and data
- Expand scope overtime

Project End-to-End Process



*Pilot Study

- Infrastructure, supported docs/macros/utility/tools, mapping spec, data programming, automate process.
- Estimate timelines and resources needs to complete one study.
- Standard implementation doc, issue tracking, escalation, resolution and decisions making process.

Case Studies

Company A (2019)

Purpose: FDA Submission

Scope of Work:

- Legacy data conversion
- Ongoing studies standardization
- Safety Data Mart

1st & 2nd Phase: 10/13 studies

Pilot study: 2-3 weeks

Timelines: 4 months

Outcomes:

- Data warehouse to support the internal safety data mart and full NDA ISS submission
- Set up standard process for ongoing studies to facilitate the data integration

Company B (2011)

Purpose: ARs, Publications, Regulatory Submissions and Queries Rapid Responses

Scope of Work:

- 12+ studies
- Initial setting & periodical run

Outcomes:

- Shorten deliverable timelines
- Reduce resources

Company C (2014)

Purpose: meta-analysis, safety & efficacy for expanded TAs

Scope of Work:

- Initial 42 CIN studies
- Expand to total 109 studies

Company D (2021-)

Purpose: FDA NDA Submission

Scope of Work:

- 5 out of 53 legacy studies data conversion, safety data warehouse (**FDA Ok'ed**)
- Ongoing studies standardization
- ISS ADaM Data Mart

Pilot study: 3 weeks

Resources: 1 Prog & 1 Biostats

Timelines: 6 months

Outcomes:

- Standardize legacy & reproduce CSR within reason
- Facilitate legacy and ongoing study NDA ISS process
- Submission-ready data

Challenges of Implementing Data Warehouse and Data Mart Project

- **What problem are you trying to solve?** Clearly understanding the goals will help you weigh potential benefits and costs.
- **How long will it take and how much will it cost?**
 - Project will require a substantial investment of effort by experienced personnel.
 - Time and cost factors: number of studies, complexity of study designs, TAs, number of unique data domains, similarity of input data, quality of input data and documentation.
- **Do you have all the necessary documentation for all studies that are a component of the project?**

Challenges of Implementing Data Warehouse and Data Mart Project

- **What will be the impact to the traceability of analysis data that was previously created or submitted?**
 - Results will have to be re-generated to ensure that they match the original results in the CSR.
 - Documentation: LDCCP, SDSP, cSDRG, ADRG, Define.
- **Have you had a discussion with FDA about whether the work is necessary?**
- **If possible, get started as soon as possible after the end of Phase II meeting with FDA.**

Summary

- ❑ Identify clinical data standardization and integration issues, find the best solutions!
- ❑ Cost and benefit of centralized data warehouse and data mart project:
 - ❑ Short-term time, cost and resources investment, the return and benefits are greater, rewarding and impactful!
 - ❑ The ultimate goal is to facilitate & shorten the submissions and review process to bring the drug to the market.
- ❑ Find opportunities for improvement, automate process, revise the project plan overtime when incorporate new studies or components.
- ❑ **Are you ready?**



Thank You

Clio Wu

cwu@chinooktx.com

www.linkedin.com/in/clio-wu-0147a015



CHINOOK

THERAPEUTICS