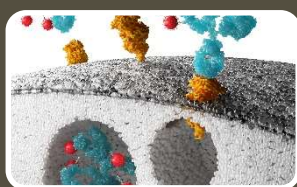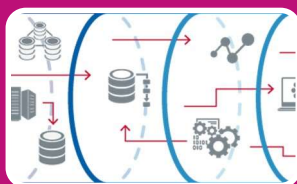# GSK R&D Data Strategy

External

Key Messages

Tech is critical to the R&D strategy to use data and AI to discovery transformational medicines

 We are trying to build deep multi-dimensional models of disease

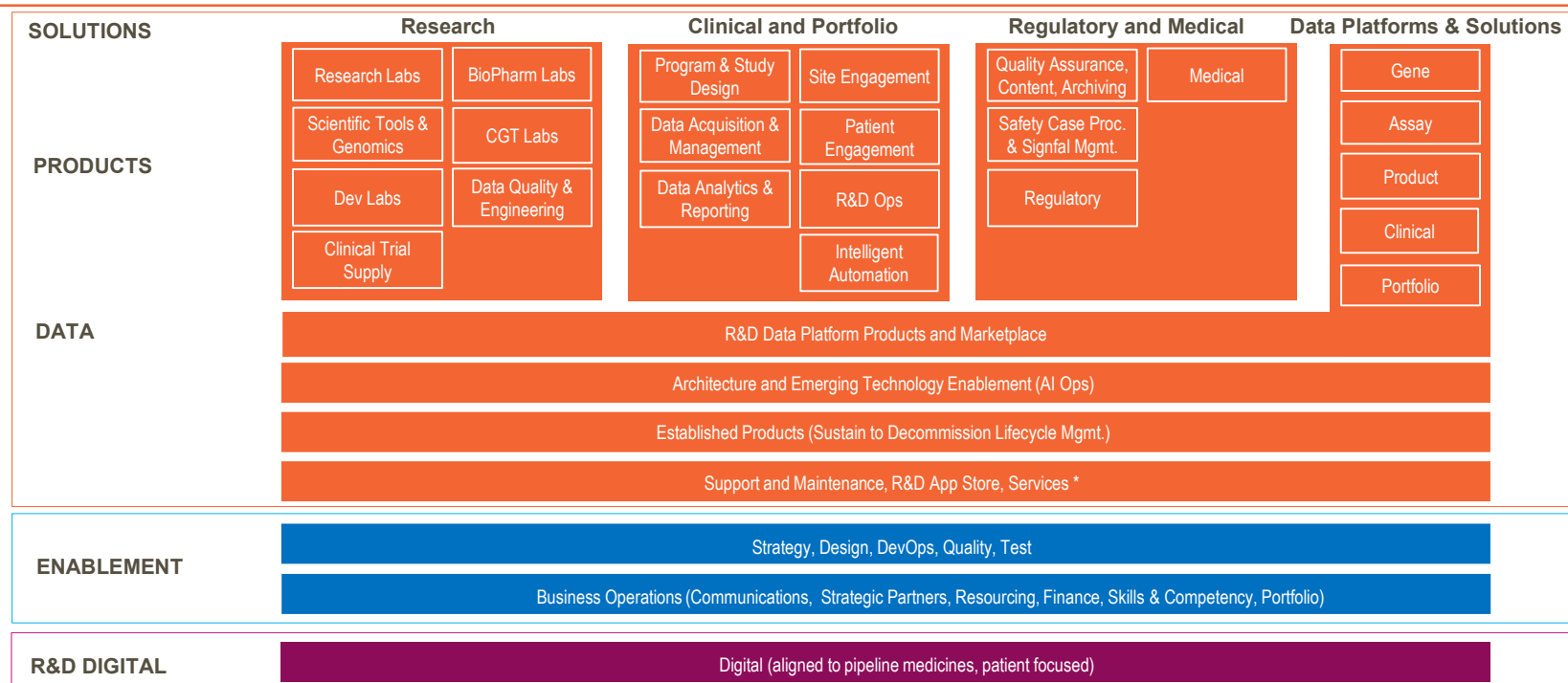 This strategy only works at scale – you need lots of different data types

 The primary challenge for Tech is to automate the end-to-end flow of data
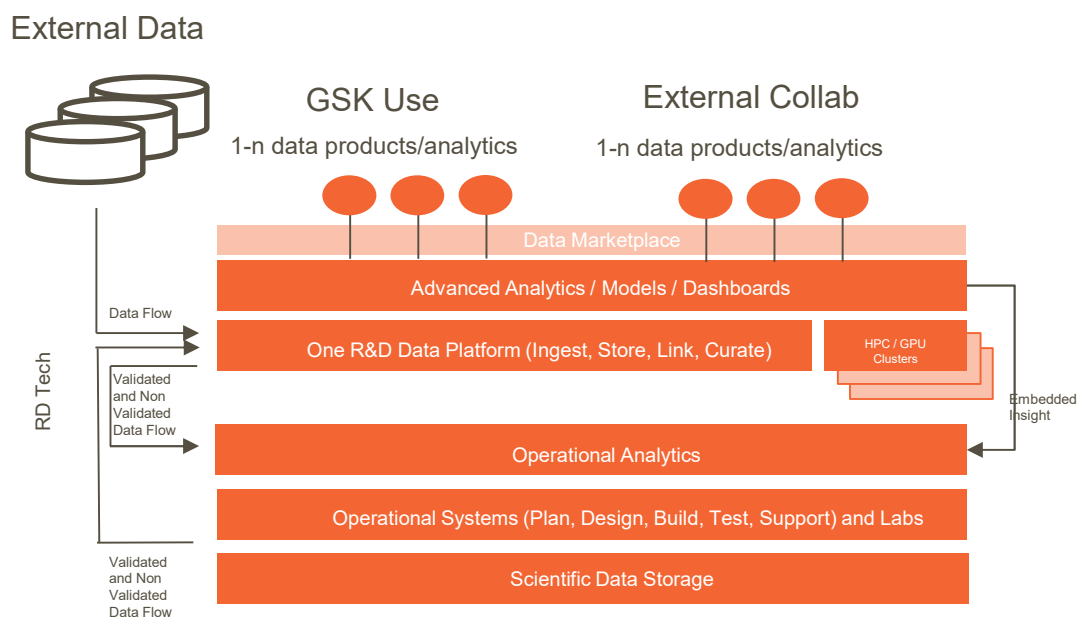
R&D Data and Tech "Product" Operating Model

Tech Solutions, Products, and Data accelerate the discovery and development of transformational medicines

gsk

SOLUTIONS	Research		Clinical and Portfolio		Regulatory and Medical		Data Platforms & Solutions
PRODUCTS	Research Labs	BioPharm Labs	Program & Study Design	Site Engagement	Quality Assurance, Content, Archiving	Medical	Gene
	Scientific Tools & Genomics	CGT Labs	Data Acquisition & Management	Patient Engagement	Safety Case Proc. & Signfal Mgmt.		Assay
	Dev Labs	Data Quality & Engineering	Data Analytics & Reporting	R&D Ops	Regulatory		Product
	Clinical Trial Supply			Intelligent Automation			Clinical
							Portfolio

DATA	R&D Data Platform Products and Marketplace
	Architecture and Emerging Technology Enablement (AI Ops)
	Established Products (Sustain to Decommission Lifecycle Mgmt.)
	Support and Maintenance, R&D App Store, Services *

ENABLEMENT	Strategy, Design, DevOps, Quality, Test
	Business Operations (Communications, Strategic Partners, Resourcing, Finance, Skills & Competency, Portfolio)

R&D DIGITAL	Digital (aligned to pipeline medicines, patient focused)

* Services include archive and financial process

R&D Data and Technology Platform

External Data

GSK Use
1-n data products/analytics

External Collab
1-n data products/analytics

Data Marketplace

Advanced Analytics / Models / Dashboards

Data Flow

RD Tech

One R&D Data Platform (Ingest, Store, Link, Curate)

HPC / GPU Clusters

Validated and Non Validated Data Flow

Embedded Insight

Operational Analytics

Operational Systems (Plan, Design, Build, Test, Support) and Labs

Validated and Non Validated Data Flow

Scientific Data Storage

How is Tech enabling this strategy?

End-to-end automated data flow is key

Operational Systems → **R&D Data Platform** → **Visualisation, Analytics, AI/ML**

Materials Management

Experiment Design

Experiment Execution

Data Processing

Data Storage

Data Model

Genome Browser

Biomarkers Analytics