

Disclaimer

Justine Rochon is employee of Takeda Pharmaceuticals

The opinions presented are solely those of the presenter and do not reflect those of Takeda

Reinventing Clinical Development: Why it's time for us to lead the AI-enabled future

Justine Rochon

PSI 2026 Conference



The Espresso Machine Lesson for AI

“

The espresso was fine, sometimes nice, mostly it was pretty bitter, sour, or quite watery. People blamed themselves, then we blamed the machine. Then we bought a second machine—same result.”

– Reuters Momentum AI NYC 2026



(Right) Questions are the Answer

“

Far better an approximate answer to the right question, which is often vague, than an exact answer to the wrong question, which can always be made precise.”

– John Wilder Tukey



Photo by Paul Halmos

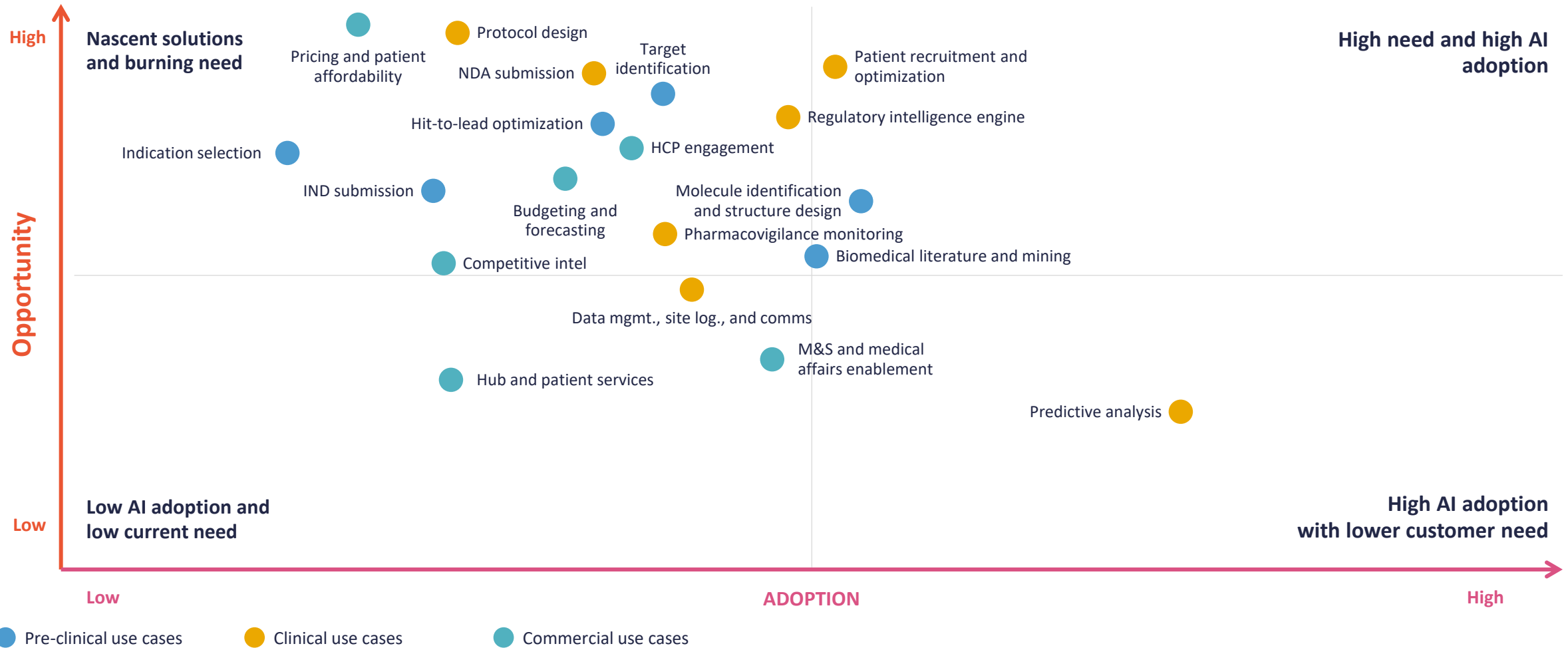


Can we afford the inefficiencies of the current system?



- Over **\$2.8 billion** to bring a single drug to market
- 9 out of 10 candidates fail before reaching patients
- Increasing legislative and regulatory pressures

Pharma Use Cases by Opportunity and Stage of AI Adoption



At Momentum AI NYC 2026, Michael Erman of Reuters interviewed Jim Swanson, CIO of J&J, during the session “Editorial Interview: Reinventing the Intelligent Enterprise,” exploring AI’s transformative role in healthcare and big pharma.



Accelerating molecular candidate identification through AI

Manufacturing process optimization through intelligent agents

Surgeon-AI collaboration in surgical procedures

AI-powered cardiac mapping for surgical efficiency

Clinical study report automation using AI agents

AI-driven clinical trial site selection and patient diversity

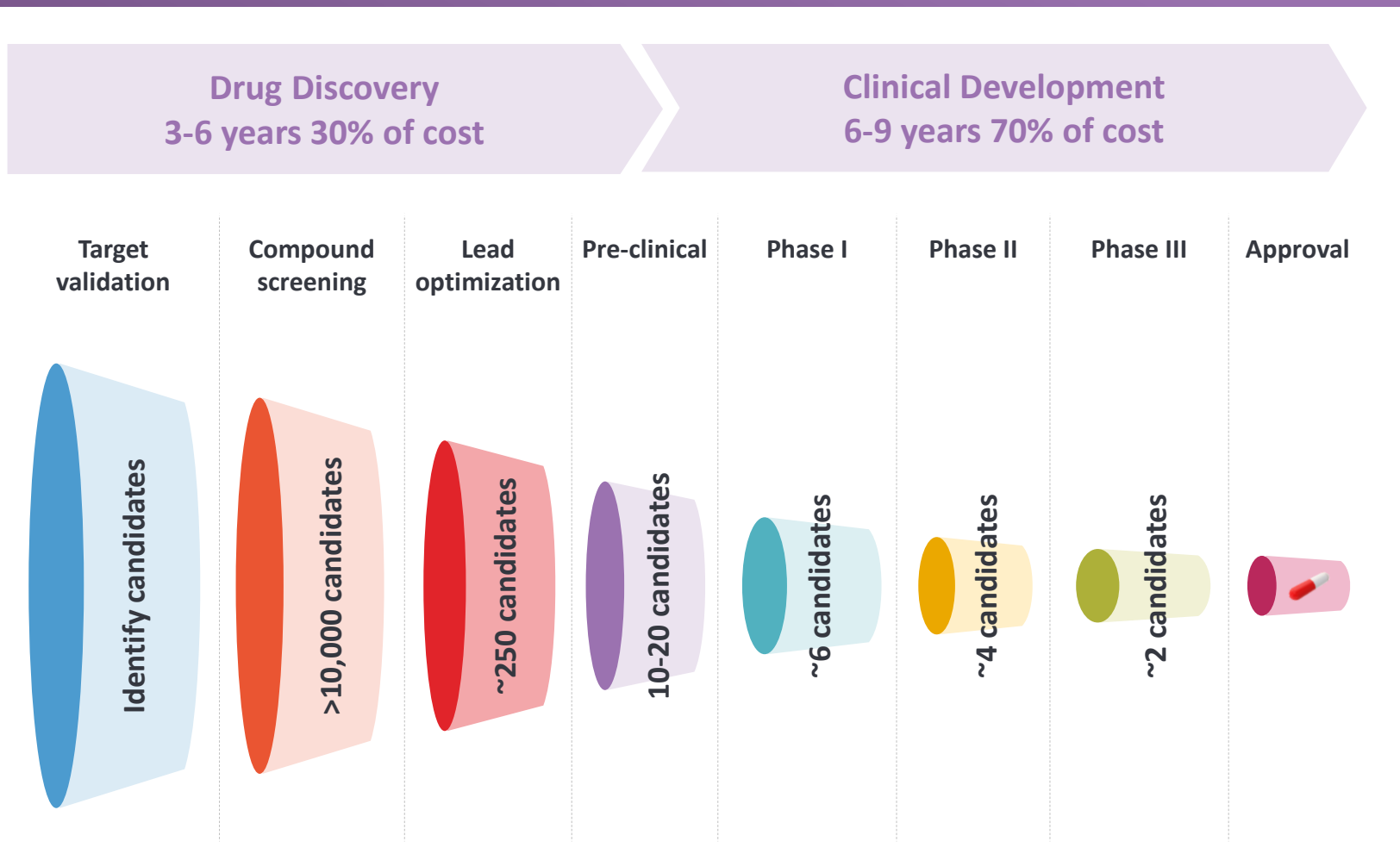
Regulatory compliance for AI-assisted drug submissions

Machine learning in compound optimization for drug discovery

Patient safety governance in AI-enabled medical devices

Building AI fluency across 140,000 employees

Drug Development is where Cost, Time, and Failure concentrate



\$ 300 billion

Annual R&D spending

12 years

Average time to market

70%

time and cost are spent on clinical trials

Clinical development concentrates most of the time, cost, and failure risk — this is where AI must deliver impact

Where AI goes next: Decision-making under Uncertainty

What AI does well today



Improves efficiency

What AI still does not solve



Trial success probability
mostly unchanged



Early-phase
uncertainty is still high

April 28, 2026 FDA issued the draft document

RFI AI-Enabled Optimization of Early-Phase Clinical Trials Pilot Program



FEDERAL REGISTER
The Daily Journal of the United States Government



Focus Areas:

- Dose selection/escalation optimization
- Real-time safety signal detection
- Go/no-go decision support with limited data
- Biomarker-driven patient selection



Kim Branson ✓ • 1st

Chief AI Officer at GSK

9mo • 🌐



So something that has begun to irritate me a little, is the constant refrain of **AI hasn't produced a drug yet,** or AI hasn't replaced X in the development process. Of course it hasn't, these things take time and very reasoned caution as we change the tech stack.

But these criticisms miss the point. It isn't about the absolute performance now, it's about the rate of change. How rapidly are the models or methods evolving, and if they have a coupled data feedback loop they will improve rapidly. Newcomen's engine didn't replace horses, but the next generation did.

Let's move the conversation to the derivative not the single point in time eval.

👍💡❤️ 661

72 comments • 13 reposts



The DISRUPT-DS Industry Roundtable:

Introduction, Overview and Latest Thinking

June 2026



DISRUPT-DS
DATA SCIENCE IN R&D ROUNDTABLE



**BOSTON
CONSULTING
GROUP**

The DISRUPT-DS Roundtable: A Forum for Data Science in Pharmaceutical R&D

DISRUPT-DS brings together senior data science leaders from major pharmaceutical companies, with the goal of

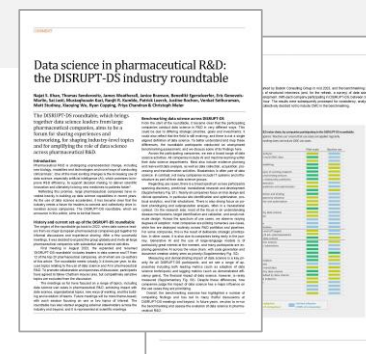
- Sharing experiences and networking
- Shaping industry-level topics
- Amplifying the role of data science across pharma R&D

Current participants: Heads of data science of major PharmaCo's

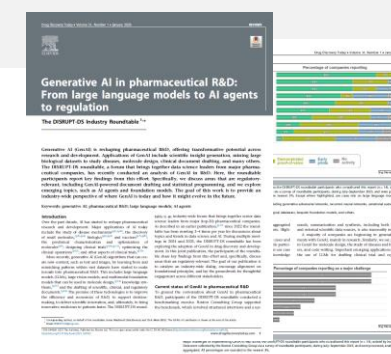
AbbVie	AstraZeneca	Bayer	BMS
Boehringer Ingelheim	EMD Serono	Gilead	Johnson & Johnson
Lilly	Merck & Co.	Novartis	Novo Nordisk
Pfizer	Roche/ Genentech	Sanofi	Takeda

Recent publications of the roundtable

State of the art of data science in pharma R&D



GenAI and AI agents in pharma R&D



Five 'Grand Challenges' for Data Science in Pharmaceutical R&D

01

Responsible & ethical use of data science & AI

02

Defining a 'backbone' of what data science in pharmaceutical R&D should include, and how to achieve and measure impact

03

Deploying data science and AI at scale

04

Ensuring the right technology infrastructure, data and other enablers, such as a FAIR data ecosystem, are in place

05

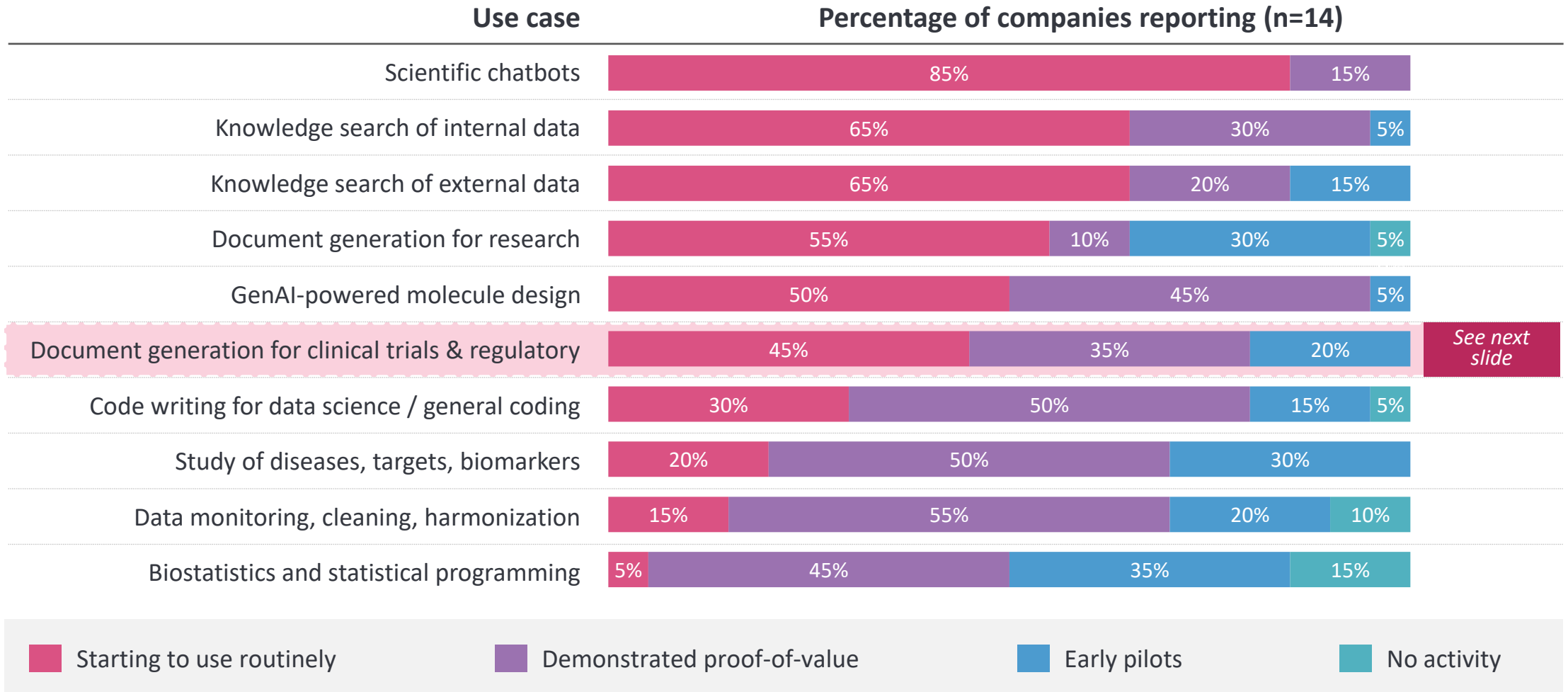
Recruiting, developing and retaining data science talent, and establishing a competency framework for data science skills

GenAI and AI Agents in Pharmaceutical R&D: Overview and State-of-the-Art today

More mature



Less mature



Document Generation for Clinical Trials & Regulatory

>90%

of large PharmaCo's have **piloted** GenAI tools for drafting R&D documents

~30%

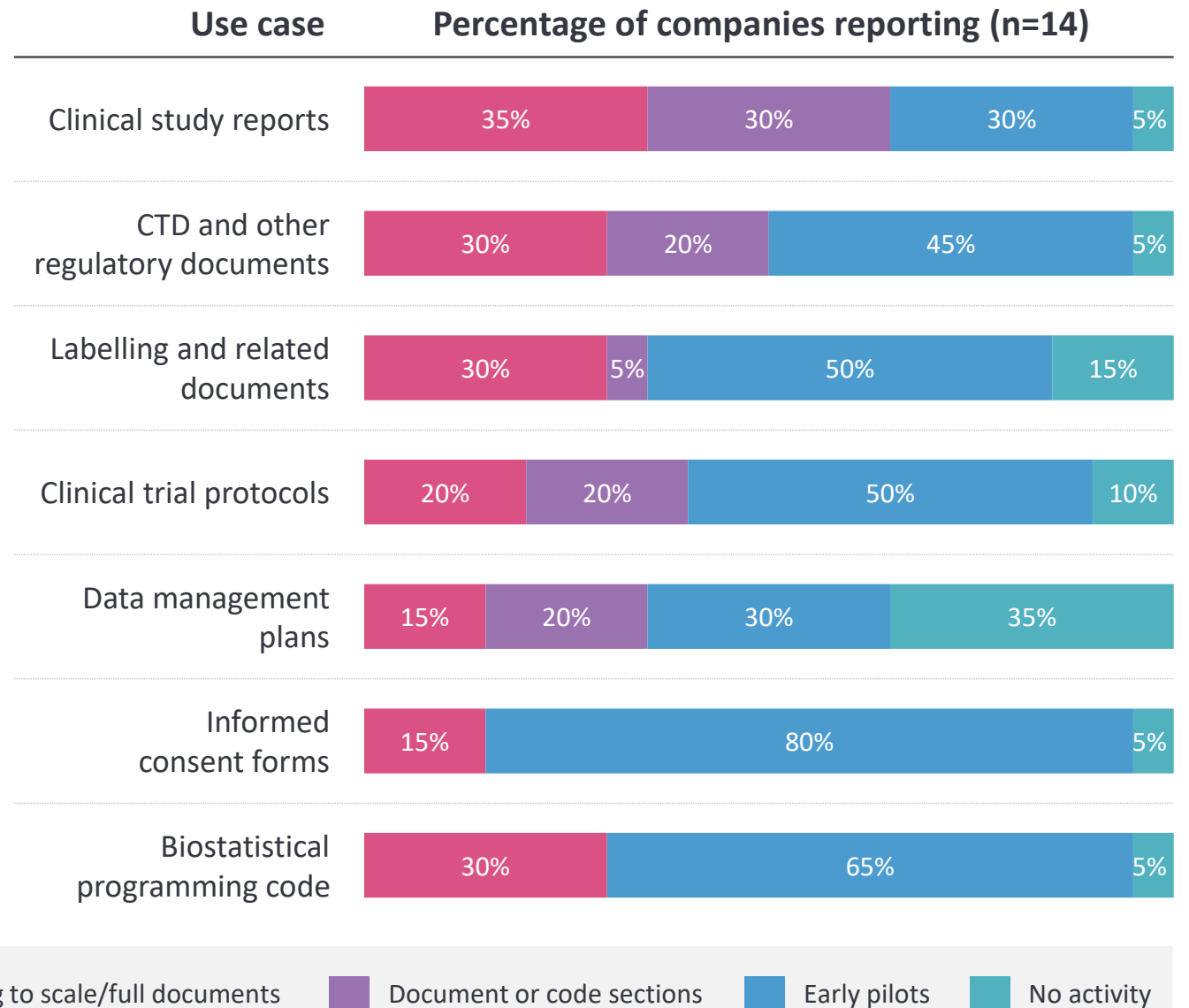
of large PharmaCo's are reporting **substantial efficiency gains** from GenAI-powered writing

20–50%

Acceleration of document writing and reviewing time, depending on document






More mature

Less mature

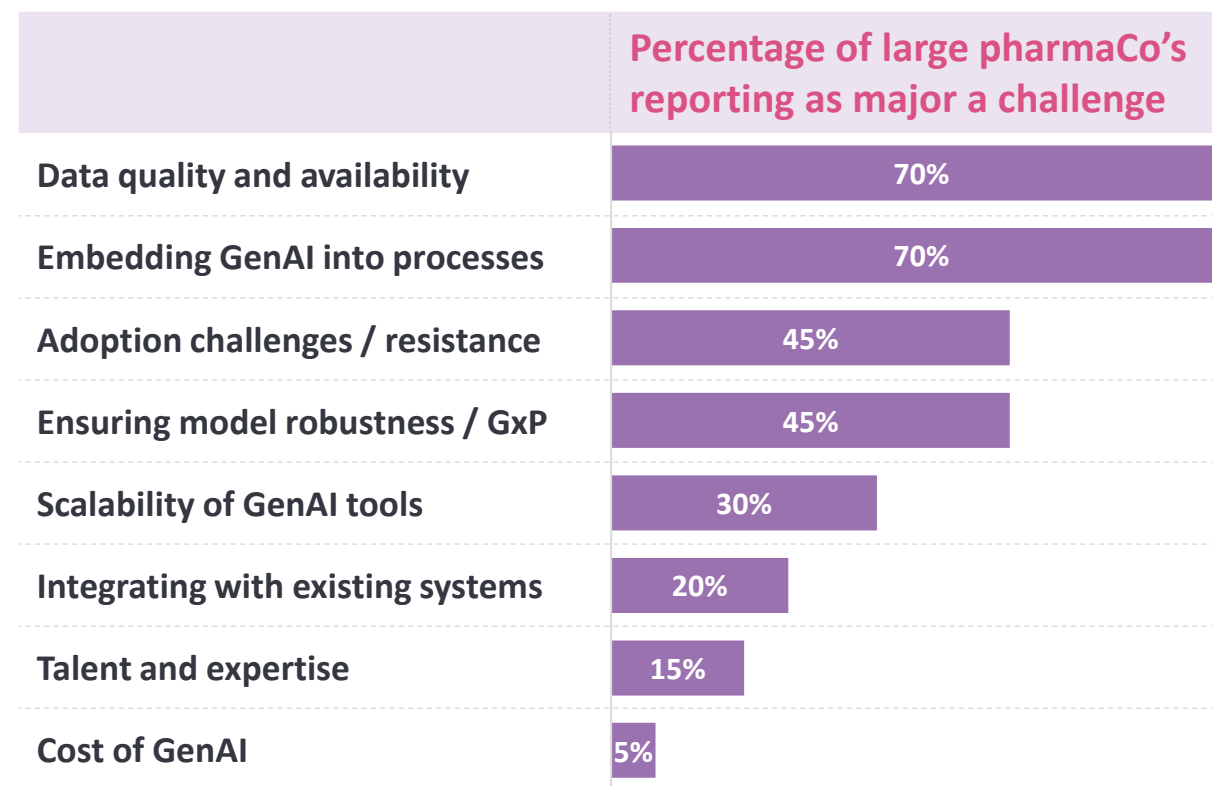


Major Challenges for GenAI in R&D

General challenges

-  Model explainability and reproducibility
-  Hallucinations
-  Biased or outdated outputs
-  Cybersecurity and data privacy
-  Intellectual property and copyright

Clinical, regulatory and medical-specific challenges



From AI Methods to Clinical-Trial Transformation: Ten Years of Research at the van der Schaar Lab Research

6 Foundations of

AI-ENABLED CLINICAL TRIALS

Ten Years of Research at the
van der Schaar Lab

01 Causal AI
Understanding treatment effects, interventions and personalized clinical decisions

02 Digital Twins
Simulating patients and trials to support better clinical decisions

03 Synthetic Data
Generating realistic healthcare data while preserving privacy and access

04 Adaptive Clinical Trials
Making clinical trials more adaptive, informative and responsive

05 AI for Pharmacology
Bridging pharmacological theory and real-world clinical data

06 Agentic AI for Clinical Trials
Supporting evidence-driven decision-making across clinical development



Revolutionizing Clinical Trials: A Manifesto for AI-Driven Transformation

This collaborative vision from academia, leaders in pharmaceuticals, and consulting, outlines how digital twins and causal inference can transform clinical trials, delivering faster, safer, and more personalized outcomes for patients

Our roadmap focuses on actionable integration within existing regulatory frameworks to revolutionize clinical research and redefine the gold standard for clinical trials using AI

Collaborators

¹Mihaela van der Schaar, ¹Richard Peck, ¹Eoin McKinney, ²Jim Weatherall, ³Stuart Bailey, ⁴Justine Rochon, ⁵Chris Anagnostopoulos, ⁶Pierre Marquet, ⁷Tony Wood, ⁷Nicky Best, ¹Harry Amad, ¹Julianna Piskorz, ¹Krzysztof Kacprzyk, ⁸Rafik Salama, ⁸Christina Gunther, ⁹Francesca Frau, ⁹Antoine Pugeat, ⁹Ramon Hernandez⁹

¹University of Cambridge, ²AstraZeneca, ³EMD Serono, ⁴Boehringer Ingelheim, ⁵QuantumBlack, ⁶University of Limoges, ⁷GSK, ⁸Accenture, ⁹Sanofi

Contents

01

The Current State of Clinical Trials

02

Key Goals of AI-Enabled Clinical Trials

03

Digital Twins: Transforming Clinical Trials for Patient-Centered Outcomes

04

Causal Inference: Transforming Clinical Trials into Engines of Insight

05

Validation Framework for AI Methods in Clinical Trials

06

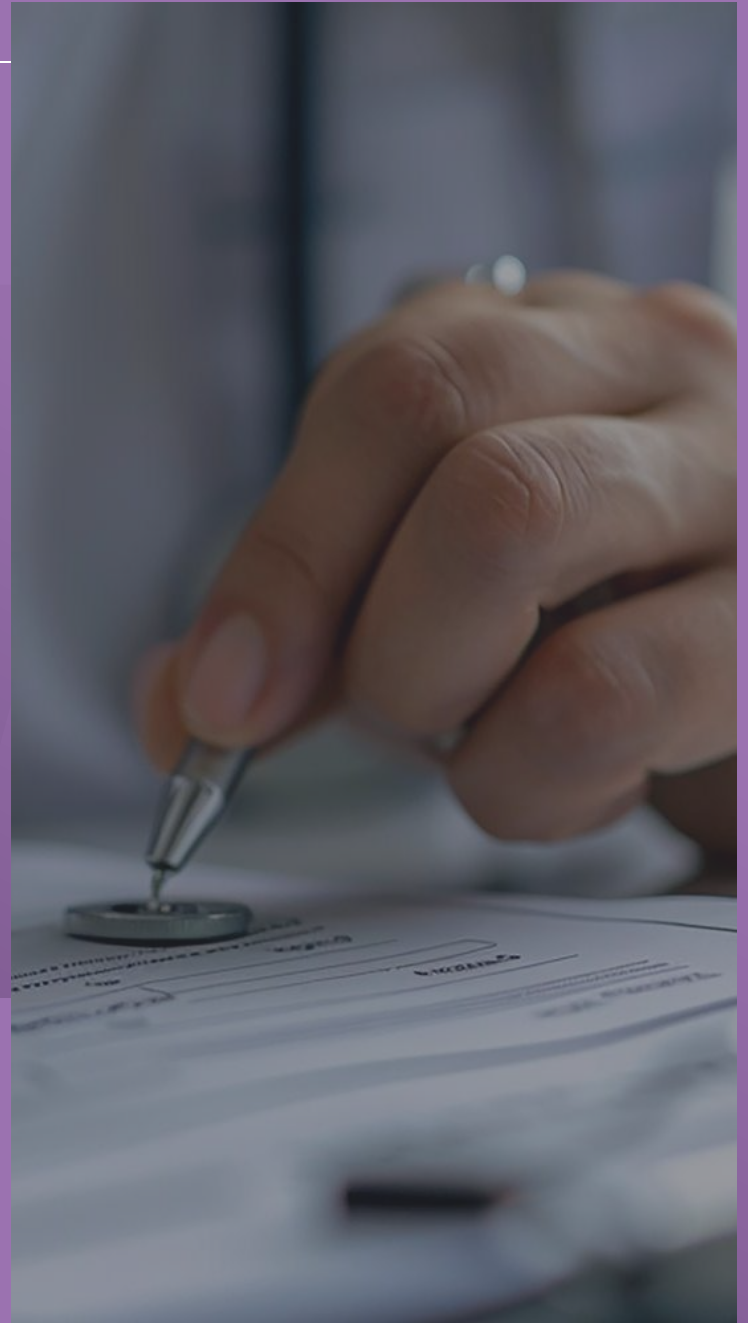
Additional Ingredients

07

Call to Action

01

The Current State of Clinical Trials



The Current State of Clinical Trials



Costly and Time-Consuming

Phase III trials often exceed \$500 million and can last several years,¹ creating significant financial burden and delaying patient access to potentially life-saving treatments



Limited Representation

Trials typically exclude over 75% of patients from testing,² although treatments are often much more widely used in diverse populations after approval³



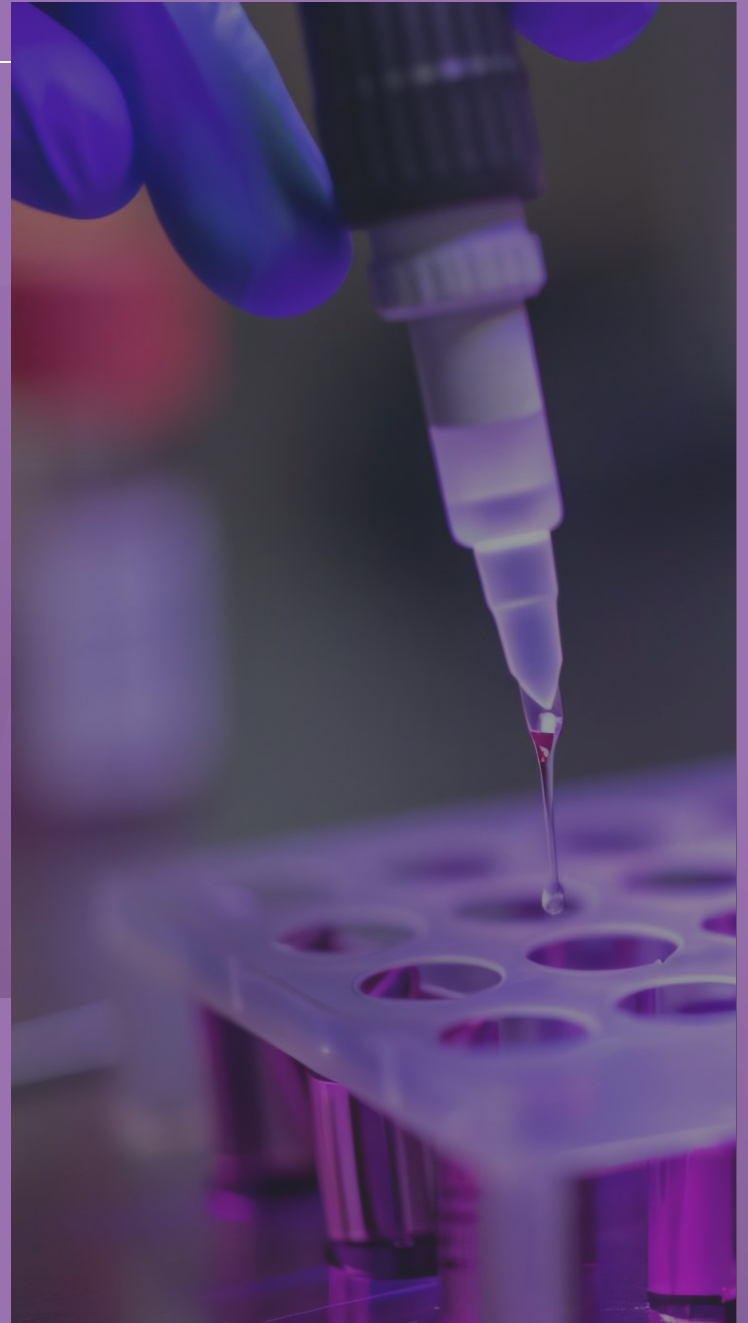
Inefficient Decision Making

Current methods often fail to identify ineffective treatments early, leading to wasted resources and delayed development of promising alternatives

1. Sertkaya et al. Costs of drug development and research and development intensity in the US, 2000-2018. *JAMA Network Open*, 7(6), 2024
2. He et al. Exclusion rates in randomized controlled trials of treatments for physical conditions: A systematic review. *Trials*, 21:1-11, 2020
3. Martin et al. Differences between clinical trials and post-marketing use. *British journal of clinical pharmacology*, 57(1):86-92, 2004

02

Key Goals of AI-Enabled Clinical Trials



Accelerating Answers with AI



Identify Responders

Causal inference identifies treatment responders with high specificity



Target Subpopulations

Enable trials to focus on those most likely to benefit



Model Patient Trajectories

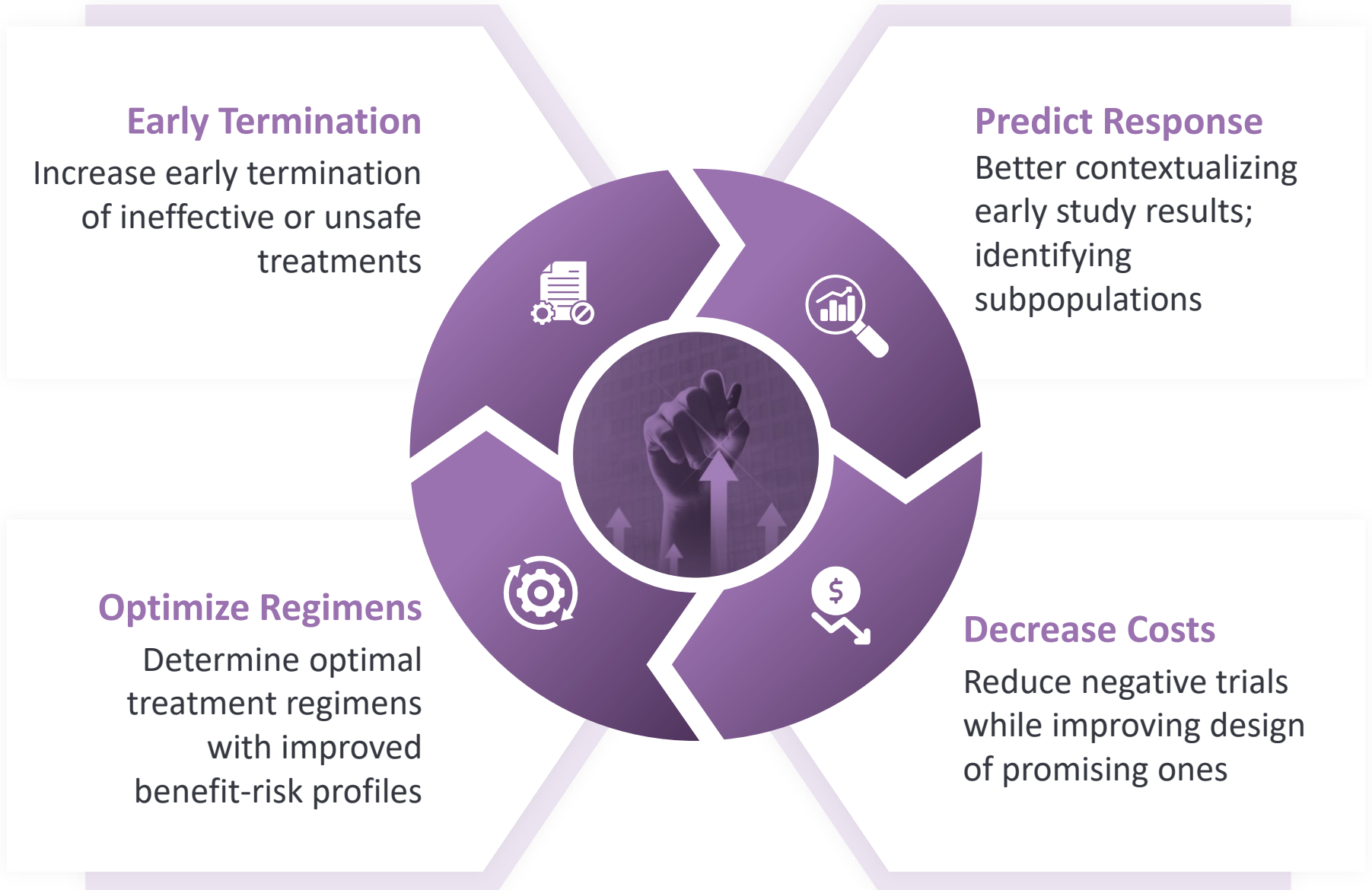
Digital twins predict safety and efficacy in real time



Faster Insights

Guide life-saving therapies to the right patients with greater urgency

Improving Probability of Success



Expanding the Questions Trials Can Answer

01



Uncovering New Effects

In silico trials uncover novel insights about treatment interactions and long-term outcomes by integrating real-world data with advanced AI methodologies

02



Combination Therapies

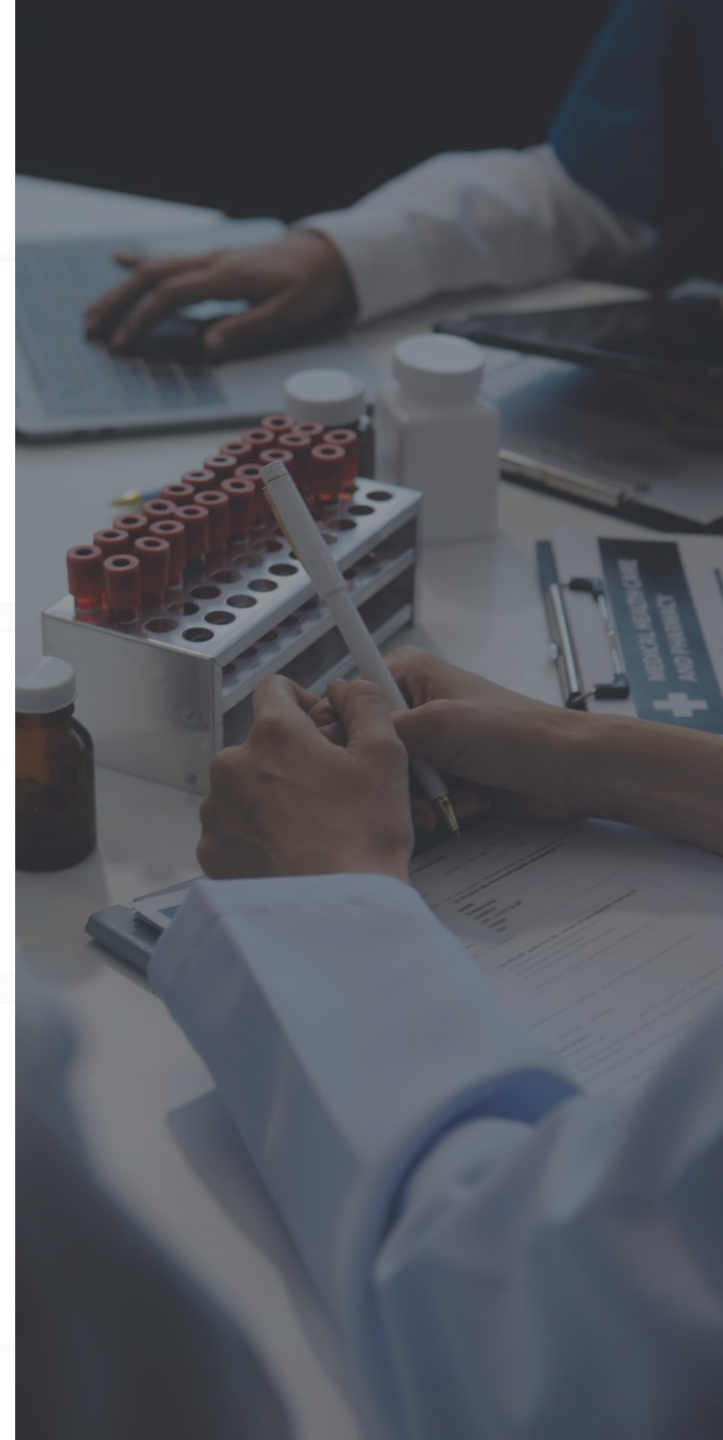
Explore how different treatments interact with each other and with lifestyle factors, providing a more holistic understanding of patient care

03



Diverse Population Responses

Simulate the impact of treatments on patients with concurrent conditions, identifying risks or synergies that conventional methods would overlook



03

Digital Twins: Transforming Clinical Trials for Patient-Centered Outcomes



Introduction to Digital Twins

Digital twins are computational models that simulate the biological and therapeutic responses of individual patients or populations. They enable real-time predictions and personalized treatment decision-making, bringing clinical trials closer to achieving the promise of precision medicine

Patient-Specific Models

Digital twins create highly personalized representations of individual patients based on their unique characteristics and medical history

AI-Driven Approach

Advanced machine learning techniques enable the creation of data-driven, adaptive digital twins that learn and evolve in real time

Counterfactual Scenarios

Twins can simulate “what-if” scenarios, testing different treatments and dosages with different patient exposure

Digital Twins Enhance Trial Diversity and Safety



Predict Safety Profiles in Phase 1–2 clinical trials

- Simulate patient reactions to various dosages and drug interaction before administration, identify risk factors
- PK/PD, leading to higher safety through intelligent treatment monitoring



Increase Participant Diversity

- Enable inclusion of patients typically excluded by classic rigid criteria, enhancing external validity, while ensuring safety through treatment adjustment



Preserve Valuable Data

- Use twins as proxies for participants who withdraw, making full use of the data already collected from such participants

Digital Twin Applications in Safety Monitoring

Cardiovascular Medicine

Integration of electrophysiological and MRI data to predict potentially fatal changes in heart rhythm after heart attacks¹

01

Metabolic Disorders

Predicting how patients with metabolic disorders might react to experimental drugs, enabling dose adjustments before administration²

02

Device Interventions

Estimating response to pacemaker implantation, facilitating targeted recruitment into trials³

03

Pharmacokinetics

Deep learning methods predicting PK responses from untested regimens to optimize dosing strategies⁴

04

MRI: Magnetic resonance imaging; PK: Pharmacokinetic

1. Arevalo et al. Nature communications, 7(1):11437, 2016

2. Bower et al. Trials, 15:1–9, 2014

3. Viola et al. Scientific reports, 13(1):8230, 2023

4. Lu et al.. iScience, 24(7), 2021

04

Causal Inference: Transforming Clinical Trials into Engines of Insight



Introduction to Causal Inference

Machine learning-based causal inference methods are revolutionizing clinical trials by ensuring the generation of actionable and reliable insights. These methods can accurately identify true treatment effects while accounting for confounding variables and selection biases



Leveraging Rich Data

Methods can utilize large-scale datasets including EHRs, biobanks, and data from failed trials to extract valuable insights



Accounting for Bias

Advanced techniques adjust for confounding variables and selection biases that can distort trial results



Beyond Correlation

Causal inference focuses on identifying true cause-and-effect relationships rather than simple correlations

Identifying Predictive Biomarkers



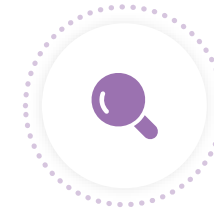
Traditional Approach

Expert-led hypotheses from domain knowledge and limited early-phase results often overlook subtle but important variables and fail to account for complex interactions



ML-Empowered Causal Inference

Focuses on Heterogeneous Treatment Effect (HTE) estimation to directly quantify how specific biomarkers influence treatment outcomes, moving beyond simplistic correlations



Systematic Screening

Leverages large-scale datasets to screen thousands of candidate variables, uncovering nonlinearities and interactions that simpler methods miss

Subgroup Analysis for Tailored Treatments

01



Traditional Analysis

Produces a single Average Treatment Effect (ATE), collapsing heterogeneous responses into one generic measure

02



Conditional Average Treatment Effect

Conditional Average Treatment Effect (CATE) characterizes a continuous spectrum of treatment responses across varying patient characteristics

03



Precision Medicine

Pinpoints groups with particularly favourable or unfavourable responses

04



Data Integration

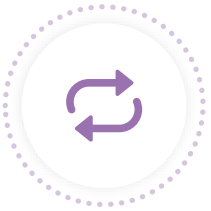
Incorporates well-curated observational data to refine subgroup definitions

Generalizing Trial Findings to Real-World Populations



Trial Environment

Controlled settings with specific inclusion criteria



Transportability Analysis

Methods to extend RCT findings to broader populations



Real-World Application

Adjusted estimates that reflect routine clinical environments and guide clinical practice



Real-World Generalization Example



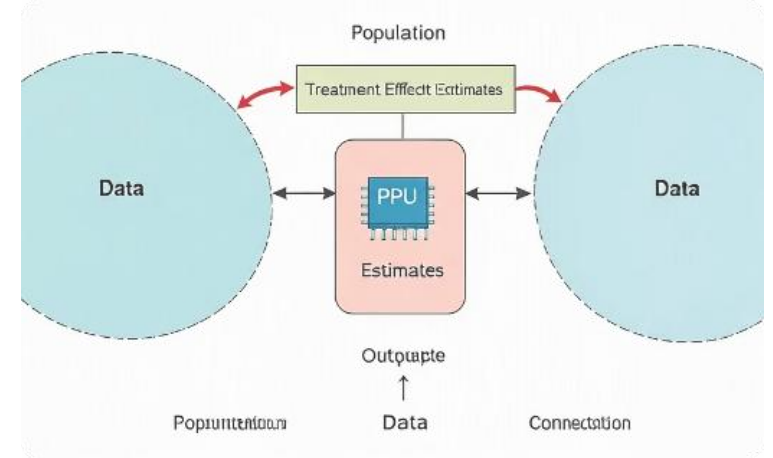
Trial Population

Novel heart failure therapy tested primarily in relatively young, low-risk patients under tightly controlled conditions



Target Population

Older, multimorbid patients in routine care settings with variable medication adherence, polypharmacy, and healthcare system differences



ML-Driven Causal Transport

Methods integrate EHR data to refine the expected benefit-risk ratio in complex real-world scenarios, guiding informed decision-making

05

Validation Framework for AI Methods in Clinical Trials



Core Components of AI Validation Framework



Define Context of Use

Specify the precise context in which the AI model will operate within the clinical trial



Establish Data Integrity

Assess whether available data accurately reflects the populations and clinical contexts of the intended application



Conduct Model Assessment

Evaluate predictive accuracy, calibration, and robustness under different conditions



Validate Interpretability

Ensure outputs are transparent and actionable for clinicians



Benchmark Against Traditional Methods

Compare AI models with existing clinical approaches to demonstrate added value

Implementation and Reporting of Validation

Integrate Validation in Real-World Settings

Testing models in their exact intended environment is crucial. For digital twins, this involves deploying alongside live trials to compare predictions against observed outcomes. For causal models, retrospective testing on historical trial data helps replicate known findings

Address Ethical and Regulatory Considerations

Ethical compliance is necessary for AI models in clinical trials. Critical examination of issues related to patient privacy and model bias is crucial, as is transparent documentation of validation procedures to gain regulatory approval

Establish Feedback Loop

AI models must evolve based on new data and insights, guided by previous validation results. Continuous feedback allows for iterative improvements in accuracy and relevance, enhancing long-term utility and reliability

Report Validation Results

Transparent reporting of validation processes and outcomes builds trust among stakeholders. Reports should address how each validation component was incorporated and what results were achieved

06

Additional Ingredients



Additional Ingredients for Success



Standardized Data Frameworks

Establish standardized, dynamic, and shareable data frameworks that ensure data from diverse sources can be harmonized and securely shared



Collaboration with Methodologists

Work closely with statistical and trial methodologists to ensure new AI technologies align with established statistical frameworks



Interdisciplinary Teamwork

Develop bidirectional knowledge exchange between AI researchers and clinical trial experts to bridge knowledge gaps

07

Call to Action



Ronny Chieng Address | Harvard Class Day 2026

Ronny Chieng Tells Harvard to 'Destroy AI' as Graduates Cheer

The comedian and *The Daily Show* host gave the keynote address for Class Day 2026.

by **Schuyler Velasco**



Ronny Chieng | PHOTOGRAPH BY STU ROSNER

*“Our generation’s upcoming battle...is going to be **people with substance** versus **people with shallow knowledge**, it’s going to be **mastery** versus **faking it**, it’s going to be people with **good taste** versus **tacky**. I trust you will put in the work necessary to be on the right side of those battles.”*

“Whatever your chosen profession is, please don’t let AI rob you of the fun part of it.”



AI Transformation Is No Longer Optional

CEOs and Boards are now setting hard mandates with specific benefit targets for AI adoption, making this transformation non-negotiable for organizations



Understanding the Four Segments of Your AI Workforce

Teams typically break down into four distinct groups: 10-15% are conscientious objectors actively resisting AI, 30% are disengaged and indifferent, 30% are excited but don't know where to start, and 10-15% are power users frustrated by the organization's slow pace



The Universal Fear: 'Am I Training My Replacement?'

The fundamental concern underlying all employee questions about AI is job security—the fear of training their own replacement. This fear emerges both early in AI exposure and resurfaces later after employees witness AI's impressive capabilities



Completely Different AI Realities

There's a dramatic perception gap between leadership and the workforce on AI adoption. While 57% of executives believe their organizations have widespread AI adoption, only 26% of managers and employees agree—revealing not just a gap, but entirely different organizational realities



The Experimentation Disconnect Is Even More Severe

The gap between executive perception and employee reality is even wider when it comes to AI experimentation. 76% of executives believe employees are encouraged to build their own AI solutions, but only 24% of the workforce feels empowered to do so

AI Washing



**AI is
Creating Jobs**



Are AI Layoffs Real?



NVIDIA CEO Jensen Huang Joins SCSP

Psychological Barriers versus Technical Barriers to AI Adoption



“

The toughest obstacles to AI adoption aren't technical – They're human. The hard part is changing the way that thousands of people work across organizations.

Two distinct moments of job displacement fear:

That fear shows up early when you're talking to somebody who hasn't quite used AI and they're not sure what it's capable of. But it shows up again later, after you have used AI and seen some of the spectacular outputs – we naturally revert to thinking about how this impacts my job.”

– Michael J Domanic, Head of AI @ Section

What does the Rise of AI mean for the Future of Statisticians?

[ASA LinkedIn Post](#)


From the campus of Harvard University, Ron Wasserstein shares a thoughtful Statistical Moment after speaking with students about one of today's most pressing questions:

What does the rise of AI mean for the future of statistics?

Across every recent visit, one theme has been constant—students are eager (and understandably curious) about how AI will shape their careers and the work of statisticians.

While predicting the future isn't simple, Wasserstein offers a clear perspective: **The foundational values of statistics—critical thinking, rigor, and sound reasoning—will only grow more important in an AI-driven world.**



The background of the slide features three wind turbines in silhouette against a vibrant sunset sky. The sky transitions from a deep purple at the top to a bright orange and yellow near the horizon, with scattered clouds catching the low light. The turbines are positioned across the frame, with one on the left, one in the center, and one on the right.

When the winds of change blow, some people build walls and others build windmills.

THANK YOU!

What will you do on next Monday to help others see that AI-enabled clinical development is statistics in action?