Scientific Learning and Persistence in Alzheimer's Drug Development*

Extended Abstract

Emma Dean Johnny Huynh Michael Nguyen-Mason
The Dartmouth Institute for Health Policy and Clinical Practice

November 2026

1 Motivation and Question

When does scientific innovation persist despite repeated failures? Drug development is inherently a trial and error process, but the high cost of failures in pharmaceutical R&D requires firms to pursue drug candidates cautiously. Clinical trial failures can provide important scientific signals, leading both developing and competing firms to terminate trials of similar drug candidates (Magazzini et al., 2012; Krieger, 2021). However, there are documented cases where drug development continues despite repeated clinical trial failures. What characteristics of the underlying science, the clinical trial failures, and the broader market allow drug development to persist despite repeated failures?

We study this question in Alzheimer's Disease (AD), a setting where drug development has been notoriously difficult. One review of the 2002–2012 research and development (R&D) pipeline estimates a 99.6 percent failure rate for Alzheimer's drug candidates (Cummings et al., 2014). Between the FDA approval of memantine in 2003 and the approval (but unsuccessful rollout) of aducanumab in 2021, no new drugs were approved to treat AD, an 18-year gap.

The difficulty of developing effective treatments for AD are particularly consequential due to the enormous social and financial burden AD places on society. An estimated 7.2 million Americans aged 65 and older are living with AD, and the cost of their health and long-term care is projected to reach $384 billion in 2025 and nearly $1 trillion by 2050 (Alzheimer's Association, 2025).

This paper asks two related questions. First, when should we expect for drug development to persist following numerous clinical trial failures? Second, what scientific and market characteristics of AD have allowed innovation to continue despite these repeated failures?

2 Background

Alzheimer's disease (AD) is multi-etiological, meaning it is believed to have multiple underlying causes, although the relative importance of different causes and their roles in disease progression are not completely understood and remain controversial. Several broad mechanisms are considered potential targets for disease-modifying treatments. The amyloid cascade hypothesis, possibly the most studied, suggests that accumulation of the amyloid-β protein fragment contributes to both development and progression of AD. An additional mechanism focuses on tau, which is a protein that

*We gratefully acknowledge funding from the NBER Coordinating Center on the Economics of Alzheimer's Disease and Alzheimer's Disease-Related Dementias (AD/AD) under NIH/NIA award U54AG090084.

accumulates abnormally in the brain in AD patients. Additional mechanisms target neuroinflammation and vascular dysfunction. This already brings variety to the mechanisms firms can target. However, even within a broad mechanism, firms can pursue different types of interventions. For example, even with drugs trying to modify amyloid-β, BACE inhibitors aim to prevent the brain from producing new amyloid-β, while amyloid antibodies aim to remove it after it has formed.

3 Model of Firm Learning

We create a simple model of firm learning from clinical trial failures to motivate our empirical analysis. First, a firm chooses a trial design x (e.g., a targeted disease phase or patient population) in intervention class j (e.g., BACE inhibitors), with available complementary technology T (e.g., a blood biomarker diagnostic test). The probability that the trial succeeds and the firm's expected payoff from running it are

$$\Pr(\text{success}) = p\, s_j\, q(x), \qquad U_j(x) = p\, s_j\, q(x)\, \Pi - c(x, T). \tag{1}$$

Here p is the firm's belief that the broad therapeutic mechanism (e.g., amyloid-β) is valid, s_j is its belief that intervention class j (e.g., BACE inhibitors) is viable, conditional on the mechanism being valid, and $q(x)$ is the probability that a trial with design x succeeds if both the mechanism and the intervention are viable. Π is the payoff if the trial succeeds, which scales with the size of the market, and $c(x, T)$ is the cost of running a trial as a function of the design characteristics x and the complementary technologies T.

Suppose that a trial in class B with design x_0 fails, and let $q_0 = q(x_0)$. Bayesian updating gives the new beliefs in the mechanism, p', and in the intervention, s'_B as:

$$\underbrace{p' = \frac{p\,(1 - s_B q_0)}{1 - p\, s_B q_0}}_{\text{updated mechanism belief}} \qquad \underbrace{s'_B = \frac{s_B\,(1 - q_0)}{1 - s_B q_0}}_{\text{updated intervention belief}} . \tag{2}$$

The firm then continues with a new trial of design x' if and only if

$$\underbrace{p'\, s'_B\, q(x')}_{\text{updated success probability}} \times \underbrace{\Pi}_{\text{payoff if successful}} \geq \underbrace{c(x', T)}_{\text{trial cost}}. \tag{3}$$

The model highlights several reasons that a trial failure may not lead to termination of a related drug. First, if q_0 is believed to be low, so that the failed design is uninformative about s_B, then the belief in the intervention barely falls. Second, if s_B is believed to be uninformative about the mechanism as a whole, then p would fall little, even as s_B falls. Third, if the market Π is large enough, as is the case with the AD market, a retrial can be deemed worthwhile even after beliefs have fallen. Fourth, if the cost $c(x', T)$ is relatively low, either because (i) external NIH funding priorities subsidize trials or (ii) the introduction of complementary technologies allows for cheaper trial designs, the continuation condition in equation (3) is easier to satisfy.

4 Data and Methods

4.1 Data

Our primary data on the AD R&D pipeline come from Clarivate's Cortellis. These data allow us to track drug programs from preclinical development through launch. Cortellis provides a rich set of trial characteristics, including the drug candidate molecule and its class, trial phase, key dates, and, where relevant, reasons for termination. It also provides information on trial design, including

enrollment criteria and measured endpoints. These data allow us to observe the complementary technologies (e.g., PET scans and blood biomarker tests) used to determine trial eligibility. We define intervention classes using Cortellis's target-based action field (e.g., BACE inhibitors).

Our sample includes only disease-modifying AD drugs, excluding drugs that treat AD symptoms. We also exclude drugs whose primary approved indication is not AD (e.g., semaglutide and statins).

4.2 Staggered differences-in-differences

We estimate the causal effect of clinical trial failures on a number of trial status and design outcomes for programs that share a broad mechanism but may not share the same intervention or patient population. Our outcomes of interest are program discontinuations, program starts, and new trials that enroll a different (earlier-stage) patient population.

To conduct this analysis we implement the staggered differences-in-differences estimator of Callaway and Sant'Anna (2021). Our primary treatment group consists of programs in the same class as the failing drug (e.g. a BACE inhibitor), while a secondary treatment group captures spillovers into other AD classes (e.g., the effect of a BACE failure on amyloid antibody programs). The event-study specification underlying our results is

$$y_{it} \;=\; \alpha_i + \gamma_t + \sum_{\substack{k=-4 \\ k\neq -1}}^{4} \beta_k \, \mathbf{1}[t - E_i = k] + \varepsilon_{it}, \tag{4}$$

where y_{it} is the outcome for program i in year t (e.g., an indicator for discontinuation), E_i is the year of the first Phase 2/3 failure in program i's class, α_i and γ_t are program and year fixed effects, and β_k traces the effect k years after the failure relative to the year before it. We estimate the β_k with the Callaway and Sant'Anna estimator, which compares each treated cohort to not-yet-treated and never-treated control programs and then aggregates the group-time effects by event time. These are the coefficients plotted in Figure 1.

For our control group we take the universe of drugs and clinical trials in Cortellis for non-AD neurodegenerative diseases. We exclude drug programs that use the same intervention class as an AD disease-modifying drug to avoid potential spillover effects in the control group.

5 Results

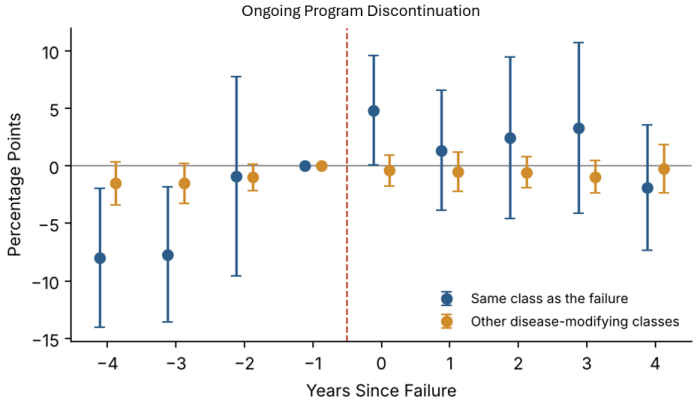
5.1 Event-study results

Figure 1 presents the results of our event studies for our three outcomes. Panel (a) shows the results of the event study regressing clinical trial failures in a class on discontinuation of ongoing programs in (i) the same class and (ii) other AD classes. While results are noisy, we do not find statistically significant program discontinuations either for same-class programs or for other-class programs.

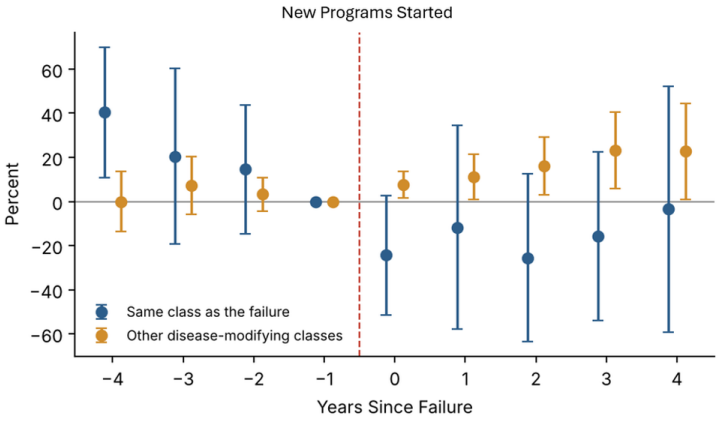
Panel (b) shows the results of the event study regressing clinical trial failures on program starts. While we find no statistically significant effect of failures on program starts in the failing class, we do find a modest and statistically significant *increase* in program starts in other AD classes. Around the third year after a class experiences a clinical trial failure, program starts in other AD classes are approximately 20 percentage points higher. The time delay makes sense given the long time horizon required to plan and start most clinical trials.

Finally, Panel (c) shows the results of the event study regressing clinical trial failures on new trials that switch to an early-stage enrollee population. We find a statistically significant increase in early-stage enrollment for new trials in the same class as the failure, of approximately 18 percentage points in the third year after the failure. This is in line with the same intuition, where it takes a

(a) Ongoing program discontinuation



(b) New programs started



(c) New trials enrolling early-stage patients

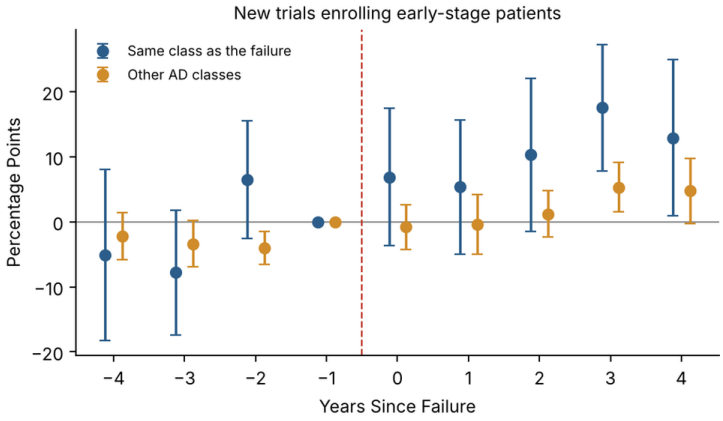


Figure 1: Event-study estimates of the effect of a clinical trial failure

fixed start-up time between a failure and the initiation of a new trial within the same program. We note that these results are driven by the earlier years of our sample. In later periods, nearly all trials enrolled patients with early-stage AD, suggesting broader learning over time.

5.2 Descriptives on complementarities in technologies

One additional hypothesis of our model is that firms are more likely to continue drug development after a failure when the cost of running subsequent trials is lower. Complementary technologies can reduce these costs, particularly when they make it easier to identify and enroll patients who are most likely to benefit from a treatment. This is especially relevant for AD prevention trials, which seek to identify individuals with underlying amyloid pathology before they develop symptoms. Historically, identifying these individuals has required costly or invasive testing. New blood biomarker tests can measure amyloid-β using a simple blood sample, potentially making it substantially less costly to screen patients for clinical trials. Consistent with this hypothesis, we find a dramatic shift toward trial designs requiring blood biomarker tests in 2025, when the FDA approved first such test. In the first half of 2026, there was a significant increase in trials for prevention of AD with five of six trials requiring blood biomarker tests for screening. Future work will examine the role of NIH funding in lowering AD trial costs, particularly for earlier-stage trials.

6 Conclusion

Our findings suggest that repeated clinical trial failures do not necessarily lead firms to abandon drug development. Instead, firms may adapt their development strategies, shifting investment toward other intervention classes and modifying trial designs to target different patient populations. These patterns are consistent with learning occurring at different levels, from the choice of intervention to the design of clinical trials. We also find evidence of rapid adoption of blood biomarker tests in AD prevention trials, highlighting the potential role of complementary technologies in reducing trial costs. More broadly, our findings illustrate how scientific learning, market incentives, and technological advances can sustain pharmaceutical innovation even in areas where clinical development has repeatedly failed.

Works Cited

Alzheimer's Association. *2025 Alzheimer's Disease Facts and Figures*. Alzheimer's Association, 2025, www.alz.org/alzheimers-dementia/facts-figures. Accessed 1 Nov. 2026.

Callaway, Brantly, and Pedro H. C. Sant'Anna. "Difference-in-Differences with Multiple Time Periods". *Journal of Econometrics*, vol. 225, no. 2, 2021, pp. 200–30.

Cummings, Jeffrey L., et al. "Alzheimer's Disease Drug-Development Pipeline: Few Candidates, Frequent Failures". *Alzheimer's Research & Therapy*, vol. 6, no. 4, 2014, p. 37.

Krieger, Joshua L. "Trials and Terminations: Learning from Competitors' R&D Failures". *Management Science*, vol. 67, no. 9, 2021, pp. 5525–48.

Magazzini, Laura, et al. "Learning from Failures or Failing to Learn? Lessons from Pharmaceutical R&D". *European Management Review*, vol. 9, no. 1, 2012, pp. 45–58.