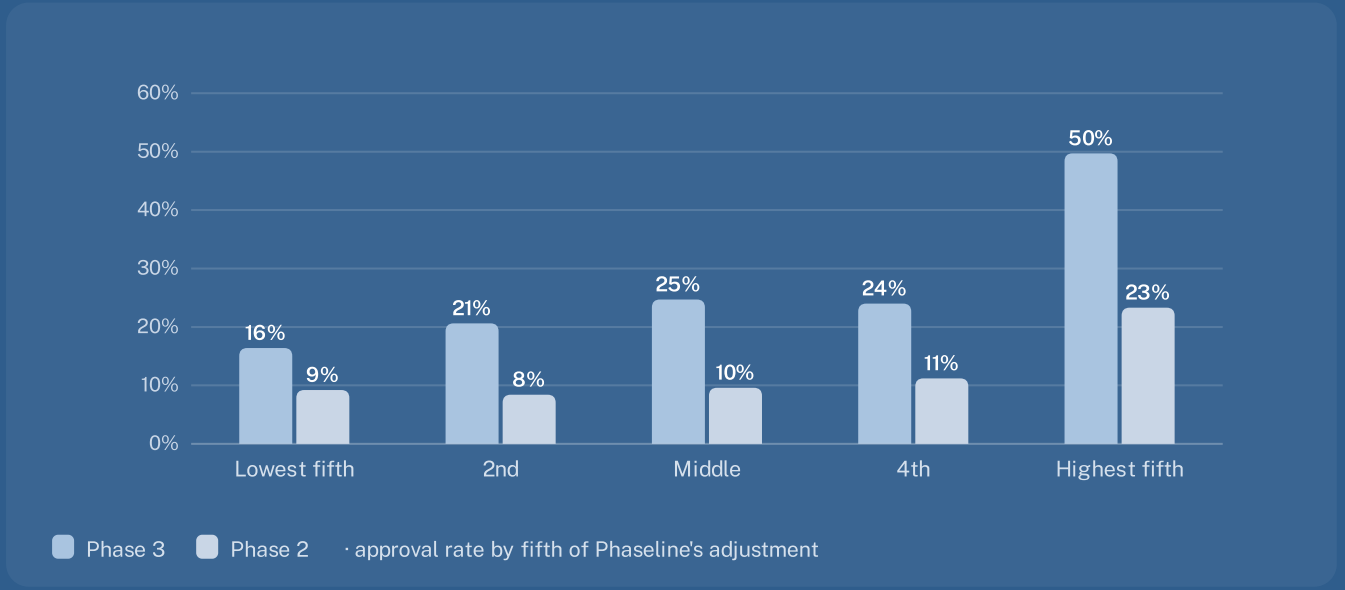
Would Phaseline have *seen it coming*?

We scored 5,693 industry drug trials that started between 2012 and 2016 with today's method, then checked which drugs the FDA approved for the disease in the trial. This is what we found, including where the method fails.



5,693
trials, 3,056 drugs

AUC 0.67
vs 0.63 for the base rate alone

50% vs 16%
approved, top vs bottom fifth, phase 3

Open
public data, one script

Summary

Would Phaseline have pointed the right way?

We took every industry drug trial in phase 2 or 3 that started between 2012 and 2016, scored each one with the method Phaseline uses today, and checked which drugs the FDA later approved for the disease in the trial.

5,693
trials of 3,056 drugs not yet approved when the trial started

19%
later FDA-approved for the trial's disease

0.67
ranking accuracy (AUC); 0.50 is a coin toss

3.0×
approval rate, strongest vs weakest fifth of Phaseline's adjustment in phase 3

What the test shows

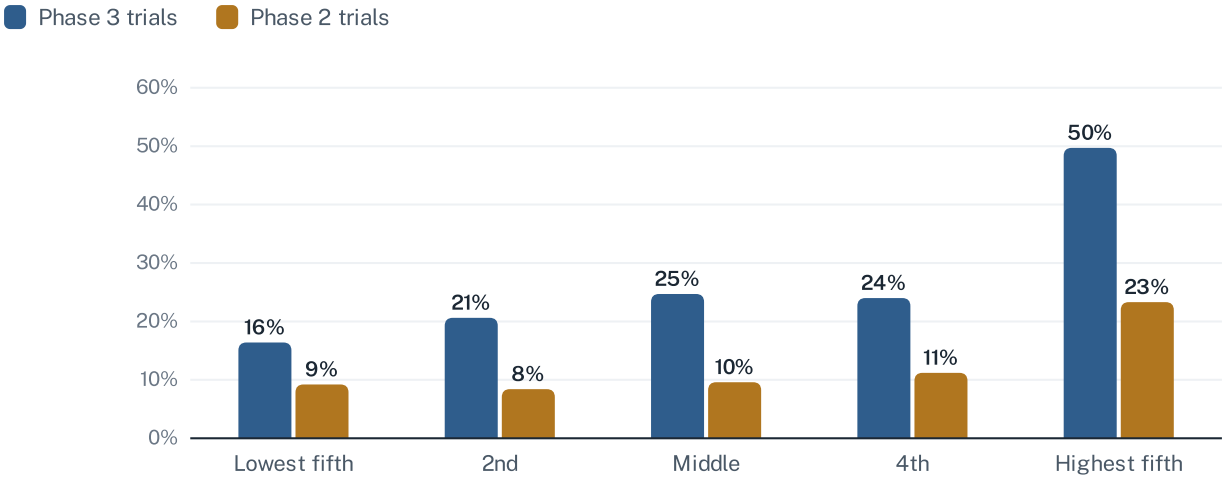
- **The ranking is real.** Across all 5,693 trials, a randomly chosen trial that led to approval had a higher Phaseline estimate than a randomly chosen one that did not in 67% of pairs (95% interval 64%–70%). The published base rate for area and phase alone reaches 63%.
- **Phaseline's own layer adds most.** Among trials that share the same base rate, the adjustment for design, drug type and genetic support separated winners from losers: in phase 3 the top fifth was approved 50% of the time and the bottom fifth 16%; in phase 2, 23% against 9%.
- **The level is somewhat optimistic.** Estimates averaged 25% while 19% of trials led to an FDA approval for their disease. Part of the gap is the test itself (FDA only, and a strict text match on the label); part is the method, and we are recalibrating it.
- **Phaseline is on the right track.** Built only from open data and published research, the method already ranks trials clearly better than the base rate alone. The test also shows exactly where to improve, and gives us a benchmark to measure every change against.

The key result

Trials Phaseline lifted were approved far more often

Each trial starts from the published success rate for its area and phase. Phaseline then multiplies it by documented factors: trial design, drug type and human-genetic support for the drug's target. Below, trials are sorted by that adjustment alone and split into fifths, so every group holds a mix of areas and the base rate cannot explain the difference.

Share of trials whose drug was later FDA-approved for the trial's disease, by fifth of Phaseline's adjustment to the base rate. Phase 3: 494 trials per fifth; phase 2: 644.

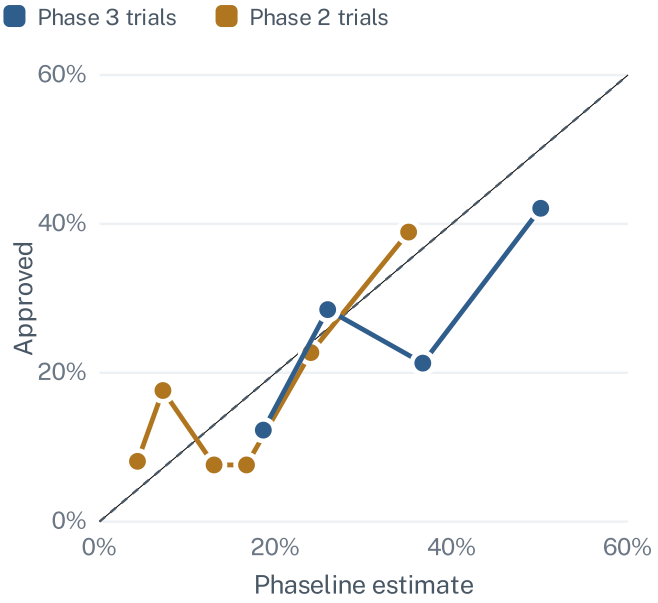


How good is the estimate?

Ranking and calibration

Measure	All	Phase 2	Phase 3
Trials	5,693	3,220	2,473
Approved for the trial's disease	19%	12%	27%
Mean Phaseline estimate	25%	15%	37%
AUC, Phaseline	0.67	0.60	0.59
AUC, base rate only	0.63	0.55	0.49
AUC within the same area and phase	0.67	0.69	0.65
Brier score (lower is better)	0.148	0.105	0.203
Brier, base rate only	0.150	0.108	0.204

AUC is the chance that a trial which led to approval scored higher than one that did not (0.5 = chance, 1 = perfect). Confidence intervals resample whole drugs, since trials of one drug are not independent.



Calibration: the dashed line is a perfect estimate. Points below it mean Phaseline was too optimistic for that group.

Did each factor point the right way?

Factor	Phase	Approved with	without	Trials	
Randomised design	2	10%	18%	2,406 / 814	✗ opposite
Blinded	2	9%	18%	2,057 / 1,163	✗ opposite
Biomarker-selected patients	2	22%	10%	590 / 2,630	✓ as expected
Genetic support for the target	2	33%	15%	318 / 1,555	✓ as expected
Antibody (vs other drug types)	2	35%	11%	198 / 3,022	✓ as expected
Randomised design	3	27%	26%	2,026 / 447	✓ as expected
Blinded	3	25%	30%	1,623 / 850	✗ opposite
Biomarker-selected patients	3	40%	24%	424 / 2,049	✓ as expected
Genetic support for the target	3	48%	34%	443 / 1,079	✓ as expected
Antibody (vs other drug types)	3	57%	24%	206 / 2,267	✓ as expected

Genetic support, antibodies and biomarker selection all pointed the way the research behind Phaseline predicts. Randomisation and blinding did not in this raw comparison: open-label, single-arm cancer trials that led to accelerated approvals outweigh them. We are reviewing both weights.

Case portfolio

Eight drugs Phaseline would have flagged

Phase 3 trials that started 2012–2016, where Phaseline's estimate was well above the base rate for the area, and the drug was later approved by the FDA for that disease. Scores are computed today from the registry record, with no knowledge of the outcome.

Evolocumab

Amgen

Phase 3 · started 2013 · NCT01764633 · Metabolic Endo

Base rate	Phaseline
31%	58%

FDA-approved 2015

PCSK9 antibody for high cholesterol. Antibody format and strong human-genetic support for PCSK9 lifted the estimate well above the base rate.

antibody · genetic support

Dupilumab

Regeneron Pharmaceuticals

Phase 3 · started 2013 · NCT01949311 · Immunology

Base rate	Phaseline
32%	52%

FDA-approved 2017

IL-4/IL-13 antibody for atopic dermatitis. Both the drug type and genetic support pointed up.

antibody · genetic support

Emicizumab

Hoffmann-La Roche

Phase 3 · started 2016 · NCT02795767 · Hematology

Base rate	Phaseline
39%	54%

FDA-approved 2017

Bispecific antibody for haemophilia A, a well-understood disease with a clear biological target.

antibody · genetic support

Fremanezumab

Teva Branded Pharmaceutical Products R&D, Inc.

Phase 3 · started 2016 · NCT02638103 · CNS Neurology

Base rate	Phaseline
27%	54%

FDA-approved 2018

CGRP antibody for migraine prevention; neurology has a low base rate, but the profile of the drug pulled it up.

antibody · genetic support

Romosozumab

Amgen

Phase 3 · started 2012 · NCT01575834 · Metabolic Endo

Base rate	Phaseline
31%	57%

FDA-approved 2019

Sclerostin antibody for osteoporosis. The SOST gene is a textbook case of genetic support.

antibody · genetic support

Risankizumab

AbbVie

Phase 3 · started 2016 · NCT02694523 · Immunology

Base rate	Phaseline
32%	53%

FDA-approved 2019

IL-23 antibody for psoriasis, an area where this mechanism class was already validated.

antibody · genetic support

Brolucizumab

Alcon Research

Phase 3 · started 2014 · NCT02307682 · Ophthalmology

Base rate	Phaseline
32%	62%

FDA-approved 2019

Anti-VEGF antibody fragment for wet macular degeneration.

antibody · genetic support

Luspatercept

Celgene

Phase 3 · started 2016 · NCT02604433 · Hematology

Base rate	Phaseline
45%	62%

FDA-approved 2019

Erythroid maturation agent for anaemia in beta-thalassaemia; haematology has one of the highest base rates.

antibody · biomarker

Where it went wrong

Misses, and what they teach

A method that is never wrong has not been tested. These are typical failures, kept in on purpose.

High estimate, no approval

Solanezumab

Eli Lilly and Company

Phase 3 · started 2013 · NCT01900665 · CNS Neurology

Base rate	Phaseline
27%	53%

Not FDA-approved for this disease

Amyloid antibody for Alzheimer's disease. Antibody format and genetic support for amyloid scored high, but the trials failed. The genetics of a disease do not guarantee that a drug hits the right stage of it.

antibody · genetic support

Bococizumab

Pfizer

Phase 3 · started 2013 · NCT01975376 · Cardiovascular

Base rate	Phaseline
30%	62%

Not FDA-approved for this disease

A second PCSK9 antibody, scored like evolocumab. Pfizer stopped it in 2016 because patients developed antibodies against the drug, a problem the registry gives no hint of.

antibody · genetic support

Roxadustat

AstraZeneca

Phase 3 · started 2014 · NCT02174627 · Hematology

Base rate	Phaseline
45%	58%

Not FDA-approved for this disease

Oral anaemia drug with a high score. It was approved elsewhere, but the FDA turned it down in 2021 over safety. Phaseline counts only FDA approvals.

genetic support · biomarker

Low estimate, approved anyway

Ibrutinib

Pharmacyclics LLC.

Phase 2 · started 2013 · NCT01744691 · Oncology

Base rate	Phaseline
5%	5%

FDA-approved 2013

BTK inhibitor in a genetically defined blood cancer. Oncology's low base rate kept the score low; the drug was approved within a year.

Ceritinib

Novartis Pharmaceuticals

Phase 2 · started 2012 · NCT01685060 · Oncology

Base rate	Phaseline
7%	7%

FDA-approved 2014

ALK inhibitor for ALK-positive lung cancer, approved on phase 2 data. The biomarker selection was in the condition, which the current method does not read.

biomarker

Targeted cancer drugs for genetically defined patients are the clearest gap: the oncology base rate is low, and the method only reads biomarkers from eligibility criteria and endpoints. Reading them from the condition and title as well is the next improvement.

What comes next

A foundation for calibration

The backtest is not a one-off. It is now part of how Phaseline is developed: every change to the method will be run against the same 5,693 trials before it reaches the site, and the results published.

Read biomarkers better
Pick up genetically defined patient groups from the condition and title, not only the eligibility criteria, so targeted cancer drugs are no longer underrated.

Recalibrate the level
Bring the average estimate in line with observed approvals, and review the weight on randomisation in phase 2.

Widen the outcome
Add approvals from the European Medicines Agency and newer start years as their outcomes become known.

Improvements will be judged on trials the method has not been tuned on, so that the track record stays honest.

Method and limits

How the test was done

Step	What we did
Cohort	All industry-sponsored interventional drug trials in phase 2 or 3 on ClinicalTrials.gov that started 2012–2016 (10,227 records). Withdrawn trials, vaccines and cell or gene therapies (not in the FDA drug database) were left out, as were trials whose drugs were all approved before the trial began.
Score	Today's Phaseline method, applied to each registry record: literature base rate by area and phase, design signal, drug type and genetic support. The area comes from the trial's conditions, or from its ICD-10 diagnosis codes when no keyword matches.
Outcome	The trial's new drug was later approved by the FDA (openFDA Drugs@FDA) *and* the trial's disease appears in the approved uses on the current FDA label (openFDA drug labels). Approvals up to 2026-10.
Reproducibility	One script, open data only. The code and the trial-level results are kept with Phaseline's source.

Limits worth knowing

- **Not fully out of sample.** The base rates and coefficients come from studies of drug development that partly overlap these years. The test shows how the method behaves on real trials; it is not a forecast made in 2014.
- **FDA only.** Drugs approved only in Europe, China or Japan count as not approved, which makes observed rates lower.
- **Text matching.** Drugs are matched by name and diseases by words on the label. Some approvals are missed and a few are matched too generously. On a looser, drug-level outcome (approved for any use) 30% of trials succeeded and the AUC was 0.61.
- **Genetics may flatter itself.** Open data about a drug's target is more complete for drugs that succeeded, which can make genetic support look stronger in hindsight.
- **Trials, not programmes.** A drug tested in many trials counts many times; intervals resample whole drugs to allow for that.

Phaseline shows population-level estimates from published research and open registries. Nothing here is investment, medical or regulatory advice. Sources: ClinicalTrials.gov (U.S. National Library of Medicine); openFDA Drugs@FDA and drug labels (U.S. FDA, public domain); Open Targets Platform (CC0); Minikel et al., Nature 2024 (CC BY 4.0).