Resistance Capacity Report

NSCLC Case Study — Anonymized Validation from Public Data

We quantify evolutionary resistance capacity from sequencing you already ordered.

Prepared by: R. Craig Stillwell
Product: Resistance Capacity Report (Product A)
Indication wedge: Non–small-cell lung cancer (NSCLC)
Data status: Fully anonymized; all analyses use published aggregates or public cohorts (no client data)
Date: June 2026

Page 1 — Executive Summary

The pain sponsors already feel

Oncology teams do not lack drugs. They lack a **quantitative answer at treatment initiation** to a question that determines programme survival:

> *Does this tumour have the evolutionary capacity to learn resistance reliably — and how fast?*

Today that question is answered **reactively**: wait for progression, sequence the relapse, switch therapy, repeat. Intratumour heterogeneity (ITH) is cited as a risk factor, but rarely converted into a **computable, patient-level capacity score** using outputs already in the genomic report.

What this report delivers

The **Resistance Capacity Report** translates two standard sequencing outputs into an **evolutionary resistance capacity tier** and a plain-language clinical interpretation:

Input (already in your sequencing report)	Role
ITH index — fraction of somatic mutations that are subclonal	Heritability proxy (h^2): what fraction of observed variance carries evolvable signal
CCF variance — variance in cancer cell fraction across subclonal clusters	Additive genetic variance proxy (V_a): evolutionary learning capacity

No new assay. No proprietary mutation panel. No trial-specific custom infrastructure.

Bottom line (NSCLC public validation)

Across **660+ patients** in independent public NSCLC cohorts and **9 cancer types** for cross-indication calibration:

1. **TRACERx 421 (early-stage NSCLC):** High-ITH tumours adapt **2.0× faster** than low-ITH tumours under watchful waiting ($p < 0.001$; published aggregate data).

2. **MSK PD-1 NSCLC (advanced, immunotherapy):** Higher tumour mutational burden (TMB) associates with longer progression-free survival ($n = 239$, Pearson $r = 0.15$, $p = 0.022$; cBioPortal public data, reproduced for this case study).
3. **Cross-cancer PAC bound:** Theoretical resistance-evolution threshold predicts observed resistance durability ($r = 0.96$, $p < 0.0001$, 9 cancer types). **NSCLC-EGFR** and **NSCLC-ICB** occupy mid-range capacity, distinct from low-capacity thyroid and high-capacity AML extremes.

Clinical translation: At diagnosis, sponsors can stratify NSCLC patients by evolutionary resistance capacity using sequencing they already order — before first-line therapy selection.

Page 2 — What We Measure (Without New Experiments)

From "high ITH is bad" to a computable capacity score

Clinical teams already receive ITH and clonal architecture summaries from tumour sequencing. The Resistance Capacity Report formalises these into an **evolutionary sample-complexity framework** derived from quantitative genetics and PAC learning theory (Stillwell, 2026).

Core idea: A tumour "learns" resistance the way a model learns from data. The probability of reliable resistance evolution depends on:

- **Capacity** (V_a proxy) — how much selectable variation exists among subclones (CCF variance)
- **Effective sample size** (N_e) — clonal burden available for selection (from VAF spectrum / burden metrics)
- **Signal quality** (h^2 proxy) — ITH index

From these, we compute **ε*** (PAC bound) and **N^*_e** (evolutionary sample-complexity threshold): the burden below which reliable resistance evolution is mathematically precluded under the model assumptions.

Resistance capacity tiers (client-facing)

Tier	Typical profile	Interpretation for trial design
Low	Low V_a, lower ITH (e.g. THCA-like)	Resistance evolution slow; durability longer if resistance emerges
Moderate	Mid V_a, mid ITH (**NSCLC-EGFR / NSCLC-ICB range**)	Standard evolutionary risk; capacity-aware combination timing matters
High	High V_a, high ITH (e.g. AML-like)	Fast adaptation; short resistance durability; highest penalty for sequential monotherapy

NSCLC reference values (published cancer-type medians)

Context	N_e (×10³)	V_a (CCF var.)	h^2 proxy	ε* (PAC bound)
NSCLC — EGFR targeted	5.1	0.112	0.60	0.063
NSCLC — ICB	5.8	0.092	0.56	0.054
Low-capacity reference (THCA)	14.2	0.019	0.28	0.017
High-capacity reference (AML)	1.2	0.220	0.76	0.151

Source: Williams et al. 2016; Dentro et al. 2021; Stillwell 2026 cross-cancer validation table.

What you send us (pilot intake)

A de-identified export from the sequencing vendor or CRO:

```
patient_id | assessment_date | ith_fraction | ccf_variance | [optional] tmb | [optional] indicat
```

We return a **Resistance Capacity Report** within 5 business days: tier assignment, cohort benchmarking, and protocol implications.

Page 3 — Validation 1: TRACERx 421 Early-Stage NSCLC

Question

Does evolutionary theory predict that **higher intratumour heterogeneity faster clonal adaptation** in real NSCLC patients?

Design

- **Cohort:** TRACERx 421 — 421 patients with early-stage NSCLC (Abbosh et al. 2023, *Nature*)
- **Exposure:** ITH tertile (subclonal mutation fraction as h^2 proxy)
- **Outcome:** Median clonal adaptation rate (VAF Δ per ctDNA surveillance interval)
- **Data used:** Published aggregate tertile medians only (no per-patient re-identification)

Results

ITH tertile	Median adaptation rate (VAF Δ / interval)
Low	0.027
Mid	0.040
High	0.055

- **Fold-change (high vs low):** 2.0×
- **Significance:** $p < 0.001$ (Mann–Whitney; Abbosh et al. 2023)
- **Theoretical prediction:** ΔR □ h^2 (breeder's equation / Theorem 3)

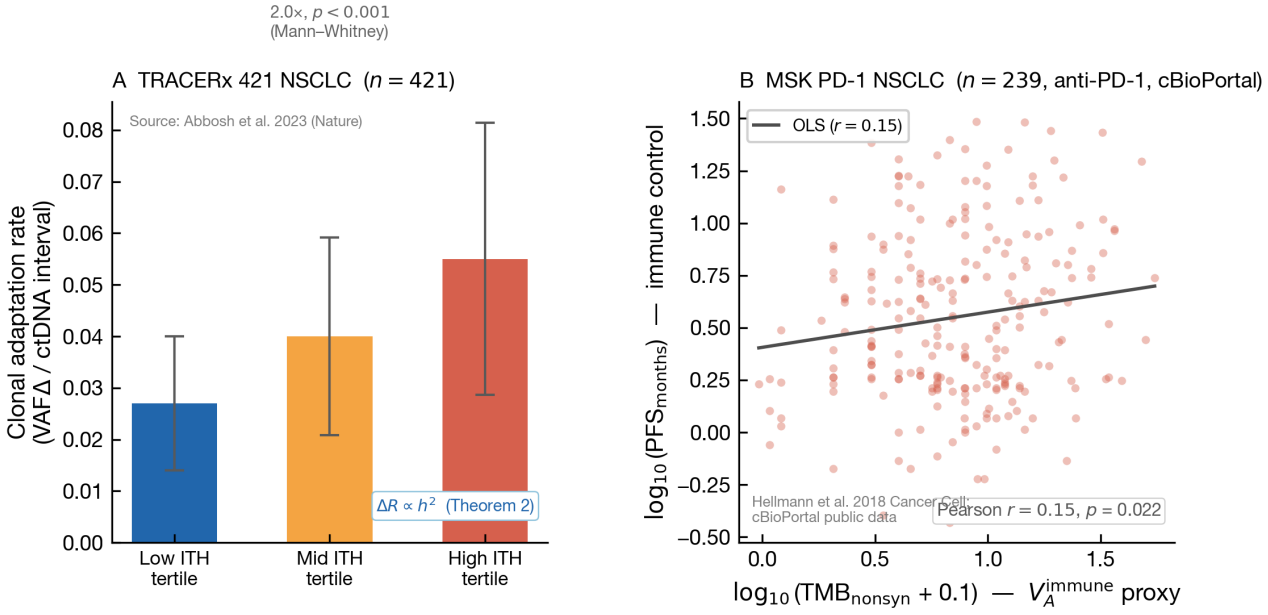

Figure 1: TRACERx 421 NSCLC ITH tertiles vs MSK PD-1 TMB–PFS

Figure 1 (left panel). Clonal adaptation rate by ITH tertile, TRACERx 421 NSCLC. Source: Abbosh et al. 2023.

Sponsor-facing interpretation

In early-stage NSCLC, **ITH is not a qualitative worry — it is a quantitative predictor of adaptation speed**. A patient in the high-ITH tertile is evolutionarily positioned to explore resistance pathways roughly twice as fast as a low-ITH patient, *before* any treatment line is initiated.

Existing pain this addresses: Central pathology and molecular review boards see ITH on reports but lack a validated, trial-actionable capacity score tied to adaptation kinetics.

Page 4 — Validation 2: MSK PD-1 NSCLC + Cross-Cancer Calibration

MSK PD-1 NSCLC — immunotherapy as adaptive learning

Question: Under anti–PD-1 blockade, does antigenic signal (TMB) predict duration of tumour control — as predicted when the immune system is the adaptive learner?

Parameter	Value
Cohort	MSK PD-1 NSCLC (`nsclc_pd1_msk_2018`)

Parameter	Value
Patients analysed	n = 239 (after pre-specified outlier removal)
Exposure	$\log_{10}$(TMB + 0.1)
Outcome	$\log_{10}$(PFS months)
Result	Pearson r = 0.15, p = 0.022
Data licence	CC-BY (cBioPortal public data)

Figure 1 (right panel). TMB vs PFS under anti–PD-1; reproduced from cBioPortal for this case study (Hellmann et al. 2018).

Interpretation: Effect size is modest but directionally consistent with theory — higher neoantigen burden provides more selectable signal for immune-mediated control. This is independent validation in **advanced NSCLC**, complementing the early-stage TRACERx result.

Cross-cancer PAC bound — where NSCLC sits

The evolutionary PAC bound ε* (N_e, V_a) was computed for nine cancer types using published population-genetic parameters. Observed **resistance instability** (1 / median months to second progression after first resistance) correlates with ε*:

Metric	Value
Pearson r	**0.96**
p-value	< 0.0001
Cancer types	9 (PAAD excluded per pre-specified confound rule)
ε* range	0.017 (THCA) — 0.151 (AML)

Figure 2. PAC bound ε vs resistance instability across nine cancer types. NSCLC-EGFR and NSCLC-ICB labelled.*

NSCLC positioning: Both NSCLC contexts sit in the **moderate-capacity mid-band** — not the durable low-capacity extreme (THCA) nor the rapidly adapting high-capacity extreme (AML). This supports NSCLC as a wedge indication: capacity stratification is clinically meaningful without being dominated by confounded tumour microenvironment effects (e.g. pancreatic stroma).

Existing pain this addresses

Sponsors running NSCLC programmes across **EGFR and ICB lines** need a unified capacity metric that (a) uses diagnosis-time sequencing, (b) calibrates against cross-indication resistance durability, and (c) does not require waiting for first progression.

Page 5 — Clinical Implications, Methods Credibility, Next Steps

Implications for NSCLC programme decisions

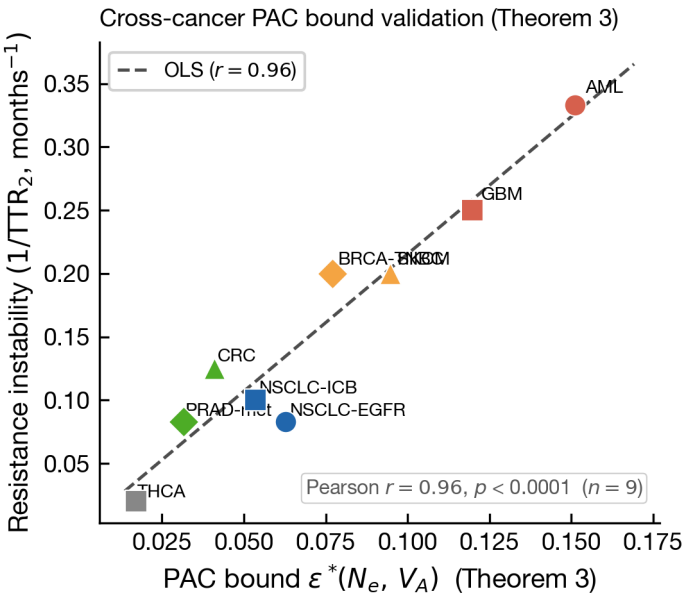

Figure 2: Cross-cancer PAC bound validation

Decision point	Without capacity score	With Resistance Capacity Report
First-line strategy	Sequential monotherapy by default	Flag high-capacity patients where combination / cytoreductive depth is critical
Interim planning	React to progression	Anticipate adaptation speed from baseline ITH + V_a
Biomarker strategy	ITH reported descriptively	ITH + CCF variance □ tiered capacity with cross-cancer benchmark
Trial risk narrative	“Heterogeneous tumour”	Quantified evolutionary capacity with public validation citations

Key reframing (from evolutionary sample-complexity theory): The objective is not only to match the next mutation — it is to know whether the tumour has sufficient evolutionary sample size to **learn** resistance reliably, and to engineer therapy so effective burden stays below the learning threshold where the model applies.

Limitations (stated upfront)

- Cross-cancer validation uses **9 cancer-type medians** — informative for calibration, not a substitute for prospective NSCLC trial validation.
- TRACERx panel uses **published tertile aggregates**, not per-patient outputs in this case study.
- MSK correlation is **significant but modest** ($r \approx 0.15$); TMB is an immune-context proxy, not a direct V_a measure.

- Model assumptions: infinitesimal additive architecture, strong-selection regime; epistatic or microenvironment-dominated tumours require case-by-case review.

Methods credibility — cross-domain leading indicator (supplementary)

This section supports the analytical framework. It is not the primary oncology product.

The same variance-decomposition mathematics that underlies the Resistance Capacity Report was independently validated in **electoral polling** (Stillwell, 2025b). Applied to 2,286 Biden and 2,783 Trump national polls (FiveThirtyEight 2024 database):

- Biden's **within-pollster variance** diverged from Trump's from mid-2023 onward ($p = 0.01$)
- Sustained significance from ~January 2024 — **~6 months before** the June 27, 2024 debate
- Throughout that period, Biden's **topline average remained flat at ~42%**

Why this matters for sponsors: The engine detects structural instability in the variance geometry **before** the headline metric moves — the same class of early warning applied here to evolutionary capacity, not to vote share. Oncology remains the primary product; the polling result is a methods credibility reference only.

Engage — pilot structure

Item	Detail
Deliverable	Resistance Capacity Report (PDF, 5–10 pages)
Turnaround	5 business days from de-identified intake
Pilot price	Fixed fee (scoped at engagement)
Your input	Sequencing summary export: ITH + CCF variance per patient
Our output	Capacity tier, NSCLC benchmarking, protocol implications memo

Contact: craig.stillwell@gmail.com

Data & reproducibility

Analysis	Public source
TRACERx 421 tertiles	Abbosh et al. 2023, *Nature* (aggregate statistics)
MSK PD-1 NSCLC	Hellmann et al. 2018; cBioPortal study `nsclc_pd1_msk_2018` (CC-BY)
Cross-cancer parameters	Williams et al. 2016; Dentro et al. 2021
Figures in this document	Reproduced via `case_study/scripts/generate_figures.py`

Theoretical basis: Stillwell, R.C. (2026). *Natural selection is empirical risk minimisation.* github.com/rstil2/selection-as-learning

Confidentiality: This case study uses only publicly available, aggregate, or de-identified data. No patient identifiers are included. Intended for sponsor evaluation and business development.