Research Square

Preprints are preliminary reports that have not undergone peer review.
They should not be considered conclusive, used to inform clinical practice,
or referenced by the media as validated information.

Diagnostic thresholds, cut-offs, and scoring controversies: a modelling study of sampling and diagnostic performance

qiaoxuan wei

358679132@qq.com

The Fourth People's Hospital of Nanning

yuting ma

The Fourth People's Hospital of Nanning

guijuan zhu

The Fourth People's Hospital of Nanning

Research Article

Keywords: diagnostic threshold, cut-off, grading, scoring system, layer distance, energy checkpoint, recovery constant, binarisation loss, SEER calibration, state-resolved readout

Posted Date: November 27th, 2026

DOI: https://doi.org/10.21203/rs.3.rs-11704622/v1

License: This work is licensed under a Creative Commons Attribution 4.0 International License.
Read Full License

Additional Declarations: No competing interests reported.

Abstract

Background.

Three controversies occupy diagnostic pathology: where a diagnostic threshold belongs, where a cut-off belongs inside a measurement range, and which scoring system measures treatment effect. Each concerns a readout several layers from the process that generates disease. We read that distance as layer distance, tested whether three models reproduce threshold loss, cut-off loss and scoring window, and used an energy checkpoint to set a common ceiling.

Methods.

We defined six layers (energy, driver, precursor, lesion, product, molecular readout) and counted layer distance d as layers between sampling channel and driver. Three PubMed blocks returned 5,640 records; a PRISMA workflow selected 75 studies. Model 1 fitted $R(d) = \exp(-\lambda d)$ to seven paired tissue–plasma comparisons. Model 2 decomposed kappa at a binary cut-off into an ordinal baseline minus binarisation loss and specimen-size loss. Model 3 fitted $R(t) = R_0 \cdot \exp(-t/\tau)$ to eight discontinuation and long-term series. Substituting the energy checkpoint E(a) of four cancers gave the age equivalent of one layer.

Results.

One layer retained 0.302–0.811 of upstream sensitivity (median 0.470), with a median decay coefficient of 0.755 per layer and a half-distance of 0.92 layers. A marker panel gained 0.08–0.45 of a layer. Binarisation cost 0.187–0.384 of kappa; specimen size cost a further 0.092. Recovery constants spanned 0.07–16.31 years (238-fold); a fixed one-year window read 0–94% of the effect. Late-phase energy decay was 0.188–0.208 per year; one layer equalled 4.01 years, and the three time scales of energy half-life, layer distance and treatment recovery (3.68, 4.01 and 4.55 years) all fell between 3.7 and 4.6 years.

Conclusions.

The three controversies share one quantity: the distance between measuring channel and process. Layer distance sets the threshold, binarisation sets the cut-off, and the recovery constant sets the scoring window. Diagnostic improvement has a ceiling at the energy end; earlier detection requires moving the sampling channel upstream, or sampling before E(a) has decayed. E(a) also decides whether a precancerous sequence exists: children hold E near its full value and carry no long precursor, whereas adults hold decayed E and pass through a precancerous stage.

1 Introduction

Diagnostic pathology currently holds three controversies.

The first concerns where a diagnostic threshold belongs: which quantity defines disease, and which value separates disease from non-disease. The second concerns where inside a measurement range a cut-off belongs: how a continuous or ordinal reading is dichotomised, and how reproducible that boundary is. The third concerns which scoring system measures treatment effect: which endpoint, at which time, judged by which rule.

The three controversies look separate and share one cause. A measuring channel always sits at some distance from the process that generates disease. A threshold on a circulating marker reports a process that lies upstream. A cut-off on a stained slide reports a reading taken from part of a lesion. A scoring endpoint measures an effect after a fixed interval, and the fraction of the effect that the interval captures depends on how fast the changed process recovers. All three quantities follow the position of the channel, and none of the three literatures treats that position as a variable.

We therefore proceed in three steps. We define six layers along the causal chain, adopt one layer of distance as the common scale, build three models, and then substitute the energy end. Model 1 addresses the threshold, Model 2 the cut-off, and Model 3 the scoring window. Three conclusions rest on three datasets and three independent variables, and one model cannot carry all three.

2 Methods

2.1 Search strategy and eligibility

We searched PubMed from inception to November 2026 in three blocks. Block 1 combined liquid-biopsy and paired-sample terms with concordance, agreement and sensitivity terms and returned 1,067 records (2015 to 2026). Block 2 combined reproducibility and observer-agreement terms with cut-off and threshold terms in pathology and immunohistochemistry and returned 1,049 records (2015 to 2026). Block 3 combined withdrawal, discontinuation and deprescribing terms with relapse and remission terms and returned 3,549 records (2010 to 2026). The three blocks yielded 5,640 unique records. Figure 1 gives the flow.

Figure 1 Identification and selection of studies.

Publication-type screening removed 103 records that were case reports, comments, editorials or letters. Title screening removed 3,392 records whose titles carried no layer, cut-off or persistence term. Abstract screening removed 2,070 records that reported a single measurement without an internal contrast, lacked a paired channel, lacked two time points, or duplicated an existing cohort. Seventy-five studies entered the models.

A study entered the analysis when it met one of three criteria. First, the same target was measured at two sampling channels within one study or one pooled analysis. Second, kappa was reported for the

same reading at two levels of aggregation, ordinal versus binary, or at two specimen sizes. Third, effect size was reported at two or more time points after an intervention started or stopped.

We excluded prediction-only data, single-layer accuracy studies without an internal contrast, case reports and narrative reviews. Tables 1 to 7 report every extracted value with its source, and any reader can check each number against the cited paper.

2.2 Six layers and their criteria

We describe disease as a chain that runs from the energy layer to the molecular readout layer. The energy layer is the supply end of the organ’s power module, and its efficiency declines with age. The driver layer is the power module, and its output sets the direction of the chain. The precursor layer holds the recognisable lesion before invasion. The lesion layer holds invasive disease that has breached the basement membrane. The product layer holds what the lesion releases into the interstitium and body cavities. The molecular readout layer holds free molecules in blood or other fluids.

The six layers follow from manipulation rather than from anatomy. A layer earns its place only when some manipulation moves that layer and leaves the others in place. Manipulating mitochondrial efficiency changes the output ceiling of the whole chain, which separates the energy layer from the driver. Removing a precursor prevents invasion without touching the driver, which separates the precursor from the driver. Excising an invasive lesion removes the source of a circulating marker without touching the driver, which separates the lesion from the driver. A blood draw measures the marker after dilution, clearance and binding, which separates the readout from the product. Each boundary carries a manipulation that tests it, and a boundary without such a manipulation does not enter the table.

Table 1
Six layers, operational criteria and typical readouts

Layer	Definition	Operational criterion	Typical readout
Energy	Supply end of the power module, whose efficiency declines with age	Manipulating mitochondrial efficiency changes the output ceiling of the chain	Mitochondrial efficiency E(a); membrane potential profile
Driver	The organ's power module, whose output sets the direction of the chain	Manipulating this unit changes the output of the whole chain	Membrane potential profile; hormonal or metabolic switch
Precursor	Morphologically recognisable lesion before invasion	Removing it prevents invasion	Serous tubal intraepithelial carcinoma; adenoma; Barrett oesophagus; cervical intraepithelial neoplasia
Lesion	Invasive disease	Breach of the basement membrane	Invasive carcinoma
Product	Material that the lesion releases into interstitium and cavities	Concentration tracks lesion burden and falls after removal	CA125; circulating tumour DNA; exosomes
Molecular readout	Free molecules in blood or other fluids	Measurable in fluid at a distance from its source	Plasma circulating tumour DNA; alpha-fetoprotein

Figure 2 Layers of the causal chain and three sampling channels.

Three parameters fix the position of a threshold. Layer distance d counts the layers between the sampling channel and the driver. Sample sufficiency A gives the proportion of specimens that complete the intended assay. Window width w gives the fraction of the disease course during which the precursor is detectable. Diagnostic gain is a function of d, A and w. The number of markers changes resolution inside a layer and leaves layer distance unchanged.

2.3 Specification of the three models

Model 1 addresses the threshold. Sensitivity retained at layer distance d decays exponentially:

$R(d) = S_{down} / S_{up} = \exp(-\lambda d)$

where S_{up} is the sensitivity of the upstream channel and S_{down} that of the downstream channel. λ is the decay coefficient in inverse layers, and the half-distance is $d_{½} = \ln 2/\lambda$. We estimated λ separately for each paired study at d = 1, because these are the comparisons in which the layer contrast is fixed and numerator and denominator come from the same specimens.

Model 2 addresses the cut-off. Agreement at a cut-off decomposes into an ordinal baseline and two losses:

κ_observed = κ_ordinal − Δ_bin − Δ_spec

Δ_bin is the binarisation loss, computed as the difference between the kappa of the ordinal scoring and the kappa of the binary cut-off drawn from the same slides and the same observers. Δ_spec is the specimen-size loss, computed as the difference between kappa at a core length of at least 50 mm and kappa below that length. Both terms are estimated within studies, so the reading, the observers and the disease stay constant while aggregation and specimen size vary. Sensitivity enters from the other side: when a study reports sensitivity for the same specimen at two thresholds, the difference gives the trade-off that the cut-off imposes.

Model 3 addresses the scoring window. Effect remaining at time t after the intervention starts or stops follows:

$R(t) = R_0 \cdot \exp(-t/\tau)$

τ is the recovery constant, the time scale over which the process returns to its prior state once the intervention stops holding it. We estimated τ from the ratio of two time points within each series, and the effect remaining at one and five years equals exp(− 1/τ) and exp(− 5/τ) respectively.

2.4 Parameters of the energy end

Our four modelling studies of cancer origin record mitochondrial efficiency as an energy checkpoint E(a) with a range of 0 to 1, and they calibrate that checkpoint against SEER incidence [76–79]; these were posted as preprints, and we take the anchors and slopes directly from them. Because λ and the energy decay rate k are both exponential constants, the ratio λ/k gives the number of years that one layer represents.

2.5 Statistical analysis

The three models report medians and ranges and do not pool estimates. They describe orderings, and the study-level value is the unit of analysis. Model comparison uses the within-study contrasts described above. Cross-model consistency asks whether one quantity orders the three phenomena in the same direction.

3 Results

3.1 Study set

Seventy-five publications entered the analysis. They cover ovarian, hepatic, pancreatic, prostatic, breast, lung, colorectal, gastric, gastroesophageal and cervical disease, together with antihypertensive, immunomodulatory, antidiabetic, haematological and bariatric intervention cohorts.

3.2 One layer of distance costs 19% to 70% of sensitivity

Seven paired comparisons measured the same target at two layers in the same patients or the same cohort (Table 2).

Table 2
Layer decay parameters for paired comparisons at d = 1

Comparison	Upstream sensitivity	Downstream sensitivity	R	λ (per layer)	d½ (layers)	Absolute loss (pp)
NCI-MATCH, tissue versus plasma [61]	100%	81.1%	0.811	0.209	3.31	18.9
Plasma versus tissue, mutation concordance [9]	100%	77%	0.770	0.261	2.65	23.0
Tissue versus plasma, 520-gene panel [8]	97.9%	72.2%	0.737	0.305	2.28	25.7
Plasma versus tissue, fusion concordance [9]	100%	47%	0.470	0.755	0.92	53.0
Ovarian cancer, tissue versus plasma [64]	100%	40.8%	0.408	0.896	0.77	59.2
Gastric cancer, tissue versus plasma [65]	100%	33.3%	0.333	1.100	0.63	66.7
Tissue versus plasma, EGFR mutation [7]	44.3%	13.4%	0.302	1.196	0.58	30.9

Across the seven comparisons, the retained fraction ranged from 0.302 to 0.811 with a median of 0.470, and the decay coefficient ranged from 0.209 to 1.196 per layer with a median of 0.755. At the median coefficient, the half-distance is 0.92 layers.

Figure 3 Sensitivity retained against layer distance.

Absolute loss and retained fraction measure different quantities. A marker with high upstream sensitivity loses many percentage points and retains a large fraction. The 520-gene panel loses 25.7 points and retains 0.737, whereas the paired EGFR assay loses 30.9 points and retains 0.302. Absolute loss varies with baseline sensitivity, and the retained fraction does not.

The same loss appears along the time axis. A multi-cancer early detection test reported 16.8% sensitivity at stage I and 90.1% at stage IV [6]. Surveillance ultrasound detected early hepatocellular carcinoma with 47% sensitivity and any-stage disease with 84% [11]. Alpha-fetoprotein reached 44.4% for early hepatocellular carcinoma against 83.5% for midkine [12], and AFP-L3 reached 34% for early disease [13].

These comparisons measure distance along the time axis rather than along the layer axis, and their retained fractions fall in a similar range.

3.3 Energy end: one layer equals 4.01 years

The energy checkpoint parameters of the four cancers appear in Table 3.

Table 3
Energy checkpoint anchors and decay rates in four cancers

Cancer	Baseline age (E ≈ 1.0)	Half age (E = 0.50)	Late-phase endpoint	Late decay rate k (per year)	Half-life (years)
Breast	65	75	0.076 at 85 y	0.188	3.68
Ovary	60	80	0.076 at 90 y	0.188	3.68
Liver	60	82	0.018 at 98 y	0.208	3.34
Lung	60	85	0.076 at 95 y	0.188	3.68

The four cancers share two features and differ in one. The late decay rate lies between 0.188 and 0.208 per year in all four, with a half-life of 3.3 to 3.7 years, a span of only 1.10-fold across cancers. The age at which E falls to 0.50 differs: 75 years in breast, 80 in ovary, 82 in liver and 85 in lung, a spread of ten years. Figure 4 plots the four curves.

Figure 4 Energy checkpoint E(a) in four cancers.

The ratio of the layer decay coefficient to the energy decay rate gives the age equivalent of one layer:

$\lambda/k = 0.755 \div 0.188 = 4.01$ years per layer

One layer of distance therefore equals about 4.0 years of energy ageing, and the half-distance of 0.92 layers equals 3.69 years.

The upstream constant is universal and the downstream constants are specific. The energy decay rate varies 1.10-fold across four cancers, whereas the layer decay coefficient varies 5.7-fold across seven paired comparisons. One constant governs the upstream end, and each organ carries its own set downstream.

3.4 A marker panel buys at most 0.45 of a layer

Adding markers raised the area under the curve by 0.04 to 0.16, which equals 0.08 to 0.45 layers at the median coefficient (Table 4).

Table 4
Within-layer gains and their equivalent layer distance

Comparison	Single marker	Panel	ΔAUC	Equivalent distance (layers)	Reference
Ovarian, CA125 to HE4 plus ROMA	0.87	0.91	0.04	0.08	[18]
Ovarian, CA125 to ROMA	0.86	0.94	0.08	0.17	[17]
Pancreatic, single to multi-marker	0.803	0.898	0.095	0.21	[15]
Pancreatic, CA19-9 to trefoil panel	0.66	0.82	0.16	0.41	[16]
Prostate, PSA to Prostate Health Index	0.60	0.76	0.16	0.45	[21]
Ovarian mucinous subgroup, CA125 to three-marker panel	47%	95%	sensitivity ratio 2.02	1.33	[20]

The largest panel gain equals 0.45 of a layer, and the smallest cross-layer loss equals one full layer. Five of the six within-layer comparisons stay below half a layer, and the single comparison that reaches a full layer is a mucinous subgroup in which CA125 performs poorly at baseline. A panel therefore buys at most 0.45 of a layer, whereas one layer of distance costs a full layer, and the two sit in different orders of magnitude.

3.5 Binarisation costs 0.19 to 0.38 of kappa

Three ring studies scored the same slides twice, once ordinally and once as a binary cut-off (Table 5).

Table 5
Binarisation and specimen-size losses in kappa

Reading	Ordinal kappa	Binary kappa	Loss	Reference
Ki-67, round 1	0.821 (5% steps)	0.437 (20% cut-off)	0.384	[56]
Ki-67, round 2	0.793 (5% steps)	0.479 (20% cut-off)	0.314	[56]
HER2-low	0.804 (5% increments)	0.617 (1% cut-off)	0.187	[55]
Breast core biopsy, grade	0.545 (≥ 50 mm)	0.453 (< 50 mm)	0.092	[29]

Figure 5 Binarisation and specimen-size losses in kappa.

The loss grows as the ordinal scoring becomes finer and as the cut-off falls in the range where most cases lie. Ki-67 at 20% loses 0.31 to 0.38, and the HER2-low boundary at 1% loses 0.19. The specimen-

size loss amounts to one half to one quarter of the binarisation loss. Within the range we examined, the choice of cut-off therefore dominates the choice of specimen.

A specimen comparison supplies the sensitivity side of the same trade-off. For claudin 18.2 in pancreatic cancer, biopsy and resection agreed in 92.5% of cases at a 75% staining threshold, and biopsy sensitivity for a positive tumour was only 54.6% [74]. Lowering the threshold to 20% raised biopsy sensitivity to 100%. Agreement stayed high while sensitivity was low because most tumours were negative at the higher threshold. The cut-off therefore decides how much disease the specimen can declare, and agreement alone does not report that quantity.

Two further studies show the same loss in expert hands. Expert breast pathologists scored HER2-low reproducibly only when categories were collapsed, with the highest agreement of 86% for 0 versus the rest and an improvement in kappa when 1 + and 2 + were combined [71]. Gynecologic pathologists reached a Fleiss kappa of 0.799 for PD-L1 CPS at the treatment threshold, and disagreements concentrated in tumours with low CPS against negative [72]. Both results follow the model: fewer boundaries give higher agreement, and the remaining disagreement sits at the boundary that the treatment decision uses.

3.6 The cost of a cut-off falls into two classes

The multi-indication PD-L1 study supplies a third class of result. Each indication had more than 100 observers, agreement was 94.1% at CPS ≥ 1 and 86.5% at CPS ≥ 20 in head and neck cancer [60]. Agreement fell as the cut-off moved into the tail of the distribution, which is the opposite of what a rule based on class proportion predicts. The agreement of a cut-off therefore depends on the reading operation it demands and not on the proportion of cases on either side.

The result divides cut-off losses into two classes. A cut-off drawn on an ordinal visual score loses through binarisation and specimen size, and the two equations above compute both. A cut-off that requires counting cells across two compartments loses through the counting itself, so the fall in CPS agreement arises from counting and not from the position of the cut-off. The two classes carry different costs and different remedies: the first improves by collapsing categories, and the second by standardising the counting rule.

3.7 The intervention layer determines the endpoint type

Table 6 lists fourteen interventions by layer and endpoint.

Table 6
Intervention layer and endpoint type

Layer	Intervention	Endpoint	Result	Endpoint type	Reference
Readout	Antihypertensive withdrawal, older adults	Restart of therapy	10.5%–33.3% versus 9%–15%	Relapse	[34]
Readout	Antihypertensive withdrawal, young adults	Remaining normotensive	0.51 at 180 days	Relapse	[35]
Readout	Anti-TNF withdrawal, Crohn's disease	Relapse	44%	Relapse	[37]
Readout	Anti-TNF withdrawal, ulcerative colitis	Relapse	38%	Relapse	[37]
Readout	Antidepressant discontinuation	Relapse	RR 0.52	Relapse	[38]
Readout	Tyrosine kinase inhibitor discontinuation	Treatment-free survival at 24 months	66%	Relapse	[39]
Readout	Cancer drugs approved in China	Overall survival gain	Median 4.1 months	Control	[40]
Driver	Prevention and screening, United States 1975–2020	Cancer deaths averted	4.75 of 5.94 million	Hard	[41]
Driver	Metabolic surgery	All-cause mortality	HR 0.508	Hard	[42]
Driver	Metabolic surgery versus medical therapy	Three-year diabetes remission	37.5% versus 2.6%	Remission	[44]
Driver	Semaglutide, obesity without diabetes	Major cardiovascular events	HR 0.80	Hard	[48]
Driver	Semaglutide, diabetes with chronic kidney disease	Major kidney events	HR 0.76	Hard	[49]
Driver	Tirzepatide, heart failure with preserved ejection fraction	Cardiovascular death or worsening heart failure	HR 0.62	Hard	[50]
Driver	Lifestyle or metformin, Diabetes Prevention Program	Diabetes incidence at 21 years	HR 0.76 and 0.83	Hard	[51]

All seven readout-layer interventions produced relapse-type endpoints after withdrawal, and all seven driver-layer interventions produced hard endpoints or remission. The two-sided Fisher exact test gives P = 0.000291.

3.8 Recovery constants span 238-fold

Eight cohorts reported effect size at two or more time points (Table 7).

Table 7
Recovery constants and effect remaining at fixed windows

Intervention	Layer of action	Time points	τ (years)	Effect at 1 year	Effect at 5 years	Reference
Antihypertensive withdrawal	Readout	180 d, 195 d	0.07	0%	0%	[35]
Autoimmune hepatitis, immunosuppression withdrawal	Readout	1, 3 y	2.60	68%	15%	[68]
Weight management (DiRECT)	Driver, behavioural	1, 2, 5 y	3.17	73%	21%	[45, 46]
Thiopurine withdrawal, inflammatory bowel disease	Readout	1, 2 y	4.29	79%	31%	[67]
Tyrosine kinase inhibitor discontinuation (DES-CML)	Readout	0, 2 y	4.81	81%	35%	[39]
Etanercept withdrawal, ankylosing spondylitis	Readout	1, 3 y	6.54	86%	47%	[69]
Bariatric surgery	Driver, structural	7, 12 y	13.90	93%	70%	[57, 59]
Imatinib discontinuation (KID)	Readout	1, 2 y	16.31	94%	74%	[70]

τ ranged from 0.07 to 16.31 years, a 238-fold span, with a median of 4.55 years.

Figure 6 Recovery time constants on a logarithmic scale with one-year and five-year windows.

The layer of action determines whether the endpoint is relapse or hard, and the recovery constant determines how long the effect reads positive at a fixed window. The two are not the same ordering. Imatinib discontinuation acts on a molecular readout, has a τ of 16.31 years, exceeds the 13.90 years of bariatric surgery and far exceeds the 3.17 years of behavioural weight management. A molecular readout can outlast a structural driver-layer change because the cell population that regenerates the readout has a longer time constant than behaviour does.

The two imatinib series also differ from each other. The DES-CML series gives 4.81 years and the KID series gives 16.31 years, a 3.4-fold difference, and the two differ in patient selection and in the duration of prior treatment. The recovery constant belongs to the population and the process rather than to the drug class, and one number for a drug class misstates it.

3.9 A fixed window reads a fraction of the effect

A one-year endpoint records 0% of the effect after antihypertensive withdrawal, 68% after immunosuppression withdrawal in autoimmune hepatitis, 73% after weight management, 94% after imatinib discontinuation and 93% after bariatric surgery. One endpoint applied to different interventions reads different fractions of the same quantity.

3.10 Three ends share one scale of about four years

The energy half-life is 3.68 years, one layer of distance equals 4.01 years, and the median treatment recovery constant is 4.55 years. The three belong to the energy, diagnostic and treatment ends respectively, and all fall between 3.7 and 4.6 years.

3.11 Childhood tumours lack a precancerous sequence

All four adult cancers carry a precancerous sequence: ductal carcinoma in situ in breast, serous tubal intraepithelial carcinoma in ovary, dysplastic nodules in liver and atypical adenomatous hyperplasia in lung. Childhood tumours mostly lack such a sequence [80, 82].

The energy end explains the difference. E(a) sits near 1.0 in children, the power module supplies enough output, and the three permissiveness gates of intrinsic limitation, tissue surveillance and systemic permission stay closed [81], so a mutation need not accumulate through a precursor layer and can realise directly as a clone. Childhood tumours therefore need fewer mutations, arise faster and carry no long precursor. Adults run the reverse: E(a) declines with age, the gates open, and a mutation must accumulate enough times in the precursor layer before it passes the three gates. Adult tumours therefore pass through a precancerous stage, and the length of that stage varies inversely with the decay rate of E(a).

The model yields one prediction: the earlier the onset age of a cancer type, the shorter its precursor sequence, to the point of being undetectable. Layer distance in Model 1 loses its precursor layer in childhood tumours, so the effective distance stays shorter than in the adult tumour of the same organ.

4 Discussion

The three controversies share one cause, and we give that cause one quantity through three models. The sections below set out the explanations and their implications.

4.1 The loss across a layer arises at four checkpoints

A marker must pass four checkpoints between production and measurement: release, entry into blood, clearance and binding. Each checkpoint costs a fraction, and the four together account for the loss across one layer. The 520-gene panel retains 0.737, whereas plasma EGFR retains 0.302. The difference lies in how well each marker passes these checkpoints, not in assay precision.

The explanation yields a selection rule. Within a layer, choose the marker that releases more, clears more slowly and binds less. Plasma half-life and free fraction estimate all three properties, and a large cohort is not required first.

4.2 The loss from binarisation arises from information compression

An ordinal score carries grading information, and a binary cut-off keeps only one side. Compression places cases on both sides of the cut-off into one class, and observer disagreement becomes visible at that boundary. Ki-67 at 20% loses 0.31 to 0.38, and HER2-low at 1% loses 0.19.

The result yields one rule: fewer cut-offs give higher agreement, and finer categories reproduce less well. The two cannot both be maximised, and the clinical use decides the trade-off. Treatment decisions call for few boundaries, and prognostic stratification can carry more.

4.3 The recovery constant is set by cell turnover

Persistence depends on how fast the cell population that regenerates the readout turns over. Blood pressure is regulated by short-term reflexes, and its τ is 0.07 years. BCR-ABL1 is regenerated from the leukaemic stem cell pool, and its τ reaches 16.31 years. Beta cells and adipose tissue turn over at intermediate rates, and the τ of DiRECT is 3.17 years. The three share one unit, and the differences arise from turnover rather than from drug class.

The result yields one rule: set the scoring window by τ rather than by a universal one year. Where τ is shorter than a year, a one-year window has already read the effect out. Where τ exceeds five years, a one-year window sees only a fraction.

4.4 The three models measure position

Layer distance measures how far the sampling channel lies from the process, the cut-off measures how far the reading lies from the centre of its range, and τ measures how far the reading lies from the process returning. All three measure position rather than substance, and the three therefore convert into one another: one layer equals 4.01 years, binarisation equals 0.19 to 0.38 of kappa, and the energy half-life is 3.68 years.

4.5 Division of labour with existing diagnostic models

Existing diagnostic models predict a diagnosis from features and measure discrimination within one layer. The present models measure how much discrimination a channel can deliver. The two run along

different axes: the first takes resolution inside a layer, and the second sets the ceiling between layers. Together they give the actual capability of a test.

4.6 Clinical implications

Four implications follow. First, move the threshold upstream: take the precursor layer where possible, and take the lesion layer rather than a body fluid where possible. Second, use few cut-offs: at a treatment threshold, collapse categories rather than add boundaries. Third, set the window by τ: match the scoring time point to the recovery constant rather than applying one year to every intervention. Fourth, stratify by E(a) rather than by calendar age.

The fourth implication carries a quantitative basis. At a layer distance of 1, E falling from 0.90 to 0.20 lowers sensitivity from 0.423 to 0.094, a 4.5-fold fall. A test that performs poorly in older patients does so because of the energy end and not because of the test.

5 Conclusions

Our conclusions form three tiers.

The first tier: the three controversies share one source. Threshold, cut-off and scoring all contest position, and once position is fixed all three become measurable.

The second tier: positions convert into one another. One layer equals 4.01 years of energy ageing, binarisation equals 0.19 to 0.38 of kappa, and recovery constants span 238-fold. The three carry different units and share one scale.

The third tier: improvement in diagnosis has a ceiling, and the ceiling sits at the energy end. E(a) follows age and organ and does not yield to any assay. The route to earlier detection is therefore to move the sampling channel upstream, or to sample before E(a) has decayed, rather than to measure the same layer better. E(a) also decides whether a precancerous sequence exists: children hold E near its full value and carry no long precursor, whereas adults hold decayed E and pass through a precancerous stage. The threshold controversy, at its root, is the energy-end controversy projected onto the sampling end.

Abbreviations

AFP: alpha-fetoprotein; AUC: area under the curve; CA125: cancer antigen 125; CML: chronic myeloid leukaemia; CPS: combined positive score; ctDNA: circulating tumour DNA; EGFR: epidermal growth factor receptor; HER2: human epidermal growth factor receptor 2; HR: hazard ratio; Ki-67: proliferation marker; PD-L1: programmed death-ligand 1; ROC: receiver operating characteristic; ROMA: Risk of Ovarian Malignancy Algorithm; SEER: Surveillance, Epidemiology, and End Results; TKI: tyrosine kinase inhibitor; TNF: tumour necrosis factor.

Declarations

Ethics approval and consent to participate. Not applicable. This study analyses published data and involves no human participants, human tissue or animal experiments.

Consent for publication. Not applicable.

Availability of data and materials. All data generated or analysed during this study are included in this published article. Tables 1 to 7 report the values that enter the three models, each with its source citation, and the studies cited in the text are identified in the reference list. The three models are fully specified by the equations in section 2.3 and the parameters in Tables 2, 5 and 7, and can be re-implemented from the article alone. No new primary datasets were generated, and no repository deposition applies.

Competing interests. The authors declare that they have no competing interests.

Funding. This research received no specific grant from any funding agency in the public, commercial or not-for-profit sectors.

Authors' contributions. Qiao-Xuan Wei conceived the framework, designed the three models, extracted the data and drafted the manuscript. Yu-Ting Ma and Gui-Juan Zhu contributed to the evidence review and revised the manuscript. All authors read and approved the final manuscript.

Acknowledgements. Not applicable.

References

1. Menon U, Gentry-Maharaj A, Burnell M, Singh N, Ryan A, Karpinskyj C, et al. Ovarian cancer population screening and mortality after long-term follow-up in the UK Collaborative Trial of Ovarian Cancer Screening (UKCTOCS): a randomised controlled trial. Lancet. 2021;397:2182–93. 10.1016/S0140-6736(21)00731-5. . PMID 33991479.
2. Jacobs IJ, Menon U, Ryan A, Gentry-Maharaj A, Burnell M, Kalsi JK, et al. Ovarian cancer screening and mortality in the UK Collaborative Trial of Ovarian Cancer Screening (UKCTOCS): a randomised controlled trial. Lancet. 2016;387:945–56. 10.1016/S0140-6736(15)01224-6. . PMID 26707054.
3. Menon U, Gentry-Maharaj A, Burnell M, Ryan A, Singh N, Manchanda R, et al. Tumour stage, treatment, and survival of women with high-grade serous tubo-ovarian cancer in UKCTOCS: an exploratory analysis of a randomised controlled trial. Lancet Oncol. 2023;24:1018–28. PMID 37657461.
4. Henderson JT, Webber EM, Sawaya GF. Screening for Ovarian Cancer: Updated Evidence Report and Systematic Review for the US Preventive Services Task Force. JAMA. 2018;319:595–606. 10.1001/jama.2017.21421. PMID 29450530.
5. Han CY, Lu KH, Corrigan G, Perez A, Kohring SD, Celestino J et al. Normal Risk Ovarian Screening Study: 21-Year Update. J Clin Oncol. 2024;42:1102–1109. 10.1200/JCO.23.00141. PMID 38194613.

6. Klein EA, Richards D, Cohn A, Tummala M, Lapham R, Cosgrove D et al. Clinical validation of a targeted methylation-based multi-cancer early detection test using an independent validation set. Ann Oncol. 2021;32:1167–1177. 10.1016/j.annonc.2021.05.806. PMID 34176681.
7. Shong LY, Lim JS, Lee HW, Tan KH, Lew SY, Lee SS, et al. Detection of EGFR mutations in patients with suspected lung cancer using paired tissue-plasma testing: a prospective comparative study with plasma ddPCR assay. Sci Rep. 2024;14:25701. 10.1038/s41598-024-76890-0. . PMID 39465302.
8. Wang M, Chen X, Wang Y, Li J, Zhang Y, Liu W, et al. Concordance Study of a 520-Gene Next-Generation Sequencing-Based Genomic Profiling Assay of Tissue and Plasma Samples. Mol Diagn Ther. 2022;26:309–22. 10.1007/s40291-022-00579-1. . PMID 35305253.
9. Sugimoto A, Matsumoto S, Udagawa H, Itotani R, Usui K, Sasaki T, et al. A Large-Scale Prospective Concordance Study of Plasma- and Tissue-Based Next-Generation Targeted Sequencing for Advanced Non-Small Cell Lung Cancer (LC-SCRUM-Liquid). Clin Cancer Res. 2023;29:1506–14. 10.1158/1078-0432.CCR-22-1749. . PMID 36201167.
10. Iams WT, Mackay M, Ben-Shachar R, Adalsteinsson VA, Parikh AR, Ogino S et al. Concurrent Tissue and Circulating Tumor DNA Molecular Profiling to Detect Guideline-Based Targeted Mutations in a Multicancer Cohort. JAMA Netw Open. 2024;7:e2351700. 10.1001/jamanetworkopen.2023.51700. PMID 38252441.
11. Tzartzeva K, Obi J, Rich NE, Parikh ND, Marrero JA, Yopp A et al. Surveillance Imaging and Alpha Fetoprotein for Early Detection of Hepatocellular Carcinoma in Patients With Cirrhosis: A Meta-analysis. Gastroenterology. 2018;154:1706–1718.e1. 10.1053/j.gastro.2018.01.064. PMID 29425931.
12. Lu Q, Li S, Lu M, Zhang Y, Wang L, Chen Y, et al. Comparison of diagnostic accuracy of midkine and AFP for detecting hepatocellular carcinoma: a systematic review and meta-analysis. Biosci Rep. 2020;40:BSR20192424. 10.1042/BSR20192424. . PMID 32039435.
13. Zhou JM, Wang T, Zhang KH. AFP-L3 for the diagnosis of early hepatocellular carcinoma: A meta-analysis. Med (Baltim). 2021;100:e27673. 10.1097/MD.0000000000027673. . PMID 34713864.
14. Lee PN, Forey BA, Coombs KJ. The effect of time changes in diagnosing lung cancer type on its recorded distribution, with particular reference to adenocarcinoma. Regul Toxicol Pharmacol. 2016;81:322–33. 10.1016/j.yrtph.2016.09.018. . PMID 27664320.
15. Kane LE, Mellotte GS, Conlon KC, Ryan BM. Diagnostic Accuracy of Blood-Based Biomarkers for Pancreatic Cancer: A Systematic Review and Meta-Analysis. Cancer Res Commun. 2022;2:1229–43. 10.1158/2767-9764.CRC-22-0190. . PMID 36969742.
16. Jahan R, Ganguly K, Smith LM, Atri P, Carmicheal J, Sheinin Y, et al. Trefoil factor(s) and CA19.9: A promising panel for early detection of pancreatic cancer. EBioMedicine. 2019;42:375–85. 10.1016/j.ebiom.2019.03.056. . PMID 30956167.
17. Suri A, Perumal V, Ammiraju K, Sivaraman G, Subramani R, Venkatesan R, et al. Diagnostic measures comparison for ovarian malignancy risk in Epithelial ovarian cancer patients: a meta-analysis. Sci

Rep. 2021;11:17308. 10.1038/s41598-021-96552-9. . PMID 34453074.

18. Wang J, Gao J, Yao H, Wu Z, Wang M, Sheng X. Diagnostic accuracy of serum HE4, CA125 and ROMA in patients with ovarian cancer: a meta-analysis. Tumour Biol. 2014;35:6127–38. 10.1007/s13277-014-1811-6. . PMID 24627132.
19. Young Han C, Bedia JS, Yang WL, Hawley N, Bergan L, Lokshin AE, et al. Autoantibodies, antigen-autoantibody complexes and antigens complement CA125 for early detection of ovarian cancer. Br J Cancer. 2024;130:861–8. 10.1038/s41416-023-02560-z. . PMID 38195887.
20. Su F, Lang J, Kumar A, Ng C, Hsieh M, Suchard MA, et al. Validation of Candidate Serum Ovarian Cancer Biomarkers for Early Detection. Biomark Insights. 2007;2:369–75. PMID 19662218.
21. Nordström T, Discacciati A, Bergman M, Clements M, Aly M, Annerstedt M, et al. Prostate cancer screening using a combination of risk-prediction, MRI, and targeted prostate biopsies (STHLM3-MRI): a prospective, population-based, randomised, open-label, non-inferiority trial. Lancet Oncol. 2021;22:1240–9. PMID 34391509.
22. Russo GI, Regis F, Castelli T, Favilla V, Privitera S, Giardina R et al. A Systematic Review and Meta-analysis of the Diagnostic Accuracy of Prostate Health Index and 4-Kallikrein Panel Score in Predicting Overall and High-grade Prostate Cancer. Clin Genitourin Cancer. 2017;15:429–439.e1. 10.1016/j.clgc.2016.12.022. PMID 28111174.
23. Oppong KW, Bekkali NLH, Leeds JS, Nayar MK, Darne A, Majumdar D, et al. Fork-tip needle biopsy versus fine-needle aspiration in endoscopic ultrasound-guided sampling of solid pancreatic masses: a randomized crossover study. Endoscopy. 2020;52:454–61. 10.1055/a-1114-5903. . PMID 32162287.
24. Balasubramanian I, Fleming CA, Corrigan MA, Redmond HP, Kerin MJ, Lowery AJ. Meta-analysis of the diagnostic accuracy of ultrasound-guided fine-needle aspiration and core needle biopsy in diagnosing axillary lymph node metastasis. Br J Surg. 2018;105:1244–53. 10.1002/bjs.10920. . PMID 29972239.
25. Figueiredo VR, Cardoso PFG, Jacomelli M, Menezes A, Kengen R, Baldotto C, et al. EBUS-TBNA versus surgical mediastinoscopy for mediastinal lymph node staging in potentially operable non-small cell lung cancer: a systematic review and meta-analysis. J Bras Pneumol. 2020;46:e20190221. 10.36416/1806-3756/e20190221. . PMID 33111752.
26. Harada S, Agosto-Arroyo E, Levesque JA, Alston E, Janowski EM, Coshatt GM, et al. Poor cell block adequacy rate for molecular testing improved with the addition of Diff-Quik-stained smears: Need for better cell block processing. Cancer Cytopathol. 2015;123:480–7. 10.1002/cncy.21561. . PMID 25955105.
27. McDermott SP, Piao J, Lio E, Jalink D, Chai J, Craddock KJ et al. Quantitative assessment of cell block cellularity and correlation with molecular testing adequacy in lung cancer. J Am Soc Cytopathol. 2016;5:196–202. doi:10.1016/j.jasc.2015.11.001. PMID 31042509.
28. Magnini D, Fuso L, Barlotta V, Bonifazi M, Volpe A, Gasparini S, et al. Molecular Testing in EBUS-TBNA Specimens of Lung Adenocarcinoma: A Study of Concordance between Cell Block Method

and Liquid-Based Cytology in Appraising Sample Cellularity and EGFR Mutations. Mol Diagn Ther. 2018;22:723–8. 10.1007/s40291-018-0359-3. . PMID 30276554.

29. Shanmugalingam A, Hitos K, Pathmanathan N. Concordance between core needle biopsy and surgical excision for breast cancer tumor grade and biomarkers. Breast Cancer Res Treat. 2022;193:151–9. 10.1007/s10549-022-06548-w. . PMID 35229238.
30. Knuttel FM, Menezes GL, van Diest PJ, Witkamp AJ, van den Bosch MA, Verkooijen HM. Meta-analysis of the concordance of histological grade of breast cancer between core needle biopsy and surgical excision specimen. Br J Surg. 2016;103:644–55. 10.1002/bjs.10128. . PMID 26990850.
31. Kalvala J, Parks RM, Cheung KL. Concordance between core needle biopsy and surgical excision specimens for Ki-67 in breast cancer - a systematic review of the literature. Histopathology. 2022;80:468–84. 10.1111/his.14555. . PMID 34473381.
32. Chen R, Qi Y, Zhang Y, Zhang J, Wang Y, Liu Y, et al. Diagnostic value of core needle biopsy for determining HER2 status in breast cancer, especially in the HER2-low population. Breast Cancer Res Treat. 2023;197:189–200. 10.1007/s10549-022-06781-3. . PMID 36346486.
33. Focke CM, Decker T, van Diest PJ. The reliability of histological grade in breast cancer core needle biopsies depends on biopsy size: a comparative study with subsequent surgical excisions. Histopathology. 2016;69:1047–54. 10.1111/his.13036. . PMID 27417415.
34. Reeve E, Jordan V, Thompson W, Sawan M, Todd A, Gammie TM, et al. Withdrawal of antihypertensive drugs in older people. Cochrane Database Syst Rev. 2020;6:CD012572. 10.1002/14651858.CD012572.pub2. . PMID 32519776.
35. Lee HY, Kim HM, Cho EJ, Lee JH, Kim KI, Kim CH, et al. Withdrawal of antihypertensive medication in young to middle-aged adults: a prospective, single-group, intervention study. Clin Hypertens. 2023;29:1. 10.1186/s40885-022-00226-1. . PMID 36593518.
36. Juraschek SP, Wang Z, Bing AM, Appel LJ, Miller ER, Chang AR, et al. Effects of Antihypertensive Deprescribing Strategies on Blood Pressure, Adverse Events, and Orthostatic Symptoms in Older Adults: Results From TONE. Am J Hypertens. 2022;35:337–46. 10.1093/ajh/hpac003. . PMID 34718403.
37. Gisbert JP, Marín AC, Chaparro M. The Risk of Relapse after Anti-TNF Discontinuation in Inflammatory Bowel Disease: Systematic Review and Meta-Analysis. Am J Gastroenterol. 2016;111:632–647. 10.1038/ajg.2016.54. PMID 27002797.
38. Zaccoletti D, Marsano A, Ostuzzi G, Tedeschi F, Bovo C, Gastaldon C, et al. Comparison of antidepressant deprescribing strategies in individuals with clinically remitted depression: a systematic review and network meta-analysis. Lancet Psychiatry. 2026;13:24–36. 10.1016/S2215-0366(25)00354-2. . PMID 41386898.
39. Murbach B, Fabarius A, Saussele S, Lauseker M, Hehlmann R, Pfirrmann M et al. Kinetics of BCR::ABL1 transcript levels and molecular relapse after tyrosine kinase inhibitors discontinuation in chronic myeloid leukemia patients: preliminary results from the DES-CML study. Front Oncol. 2024;14:1393191. 10.3389/fonc.2024.1393191. PMID 38779092.

40. Zhang Y, Chen P, Niu M, Wang Y, Li Y, Zhang X, et al. Overall Survival Benefits of Cancer Drugs Approved in China From 2005 to 2020. JAMA Netw Open. 2022;5:e2225973. 10.1001/jamanetworkopen.2022.25973. . PMID 35947385.
41. Goddard KAB, Feuer EJ, Mandelblatt JS, Meza R, Holford TR, Jeon J, et al. Estimation of Cancer Deaths Averted From Prevention, Screening, and Treatment Efforts, 1975–2020. JAMA Oncol. 2025;11:162–7. 10.1001/jamaoncol.2024.5552. . PMID 39636625.
42. Syn NL, Cummings DE, Wang LZ, Lin DJ, Zhao JJ, Loh M, et al. Association of metabolic-bariatric surgery with long-term survival in adults with and without diabetes: a one-stage meta-analysis of matched cohort and prospective controlled studies with 174 772 participants. Lancet. 2021;397:1830–41. 10.1016/S0140-6736(21)00591-8. . PMID 33965067.
43. Puzziferri N, Roshek TB 3rd, Mayo HG, Gallagher R, Belle SH, Livingston EH. Long-term follow-up after bariatric surgery: a systematic review. JAMA. 2014;312:934–942. 10.1001/jama.2014.10706. PMID 25182102.
44. Kirwan JP, Courcoulas AP, Patti ME, Hu B, Gaur V, Wang J, et al. Diabetes Remission in the Alliance of Randomized Trials of Medicine Versus Metabolic Surgery in Type 2 Diabetes (ARMMS-T2D). Diabetes Care. 2022;45:1574–83. 10.2337/dc21-1756. . PMID 35320365.
45. Lean MEJ, Leslie WS, Barnes AC, Brosnahan N, McCombie L, Peters C, et al. Primary care-led weight management for remission of type 2 diabetes (DiRECT): an open-label, cluster-randomised trial. Lancet. 2018;391:541–51. 10.1016/S0140-6736(17)33102-1. . PMID 29221645.
46. Lean MEJ, Leslie WS, Barnes AC, Brosnahan N, McCombie L, Peters C, et al. 5-year follow-up of the randomised Diabetes Remission Clinical Trial (DiRECT) of continued support for weight loss maintenance in the UK: an extension study. Lancet Diabetes Endocrinol. 2024;12:233–46. 10.1016/S2213-8587(24)00069-X. . PMID 38423026.
47. Kim J, Lee DY, Park SE, Kim J, Lee JH, Kim NH, et al. Lifestyle Changes and Remission in Patients With New-Onset Type 2 Diabetes: A Nationwide Cohort Study. J Korean Med Sci. 2025;40:e24. 10.3346/jkms.2025.40.e24. . PMID 39995256.
48. Lincoff AM, Brown-Frandsen K, Colhoun HM, Deanfield J, Emerson SS, Esbjerg S, et al. Semaglutide and Cardiovascular Outcomes in Obesity without Diabetes. N Engl J Med. 2023;389:2221–32. 10.1056/NEJMoa2307563. . PMID 37952131.
49. Perkovic V, Tuttle KR, Rossing P, Mahaffey KW, Mann JFE, Bakris G, et al. Effects of Semaglutide on Chronic Kidney Disease in Patients with Type 2 Diabetes. N Engl J Med. 2024;391:109–21. 10.1056/NEJMoa2403347. . PMID 38785209.
50. Packer M, Zile MR, Kramer CM, Bauersachs J, Senni M, Bulmer B, et al. Tirzepatide for Heart Failure with Preserved Ejection Fraction and Obesity. N Engl J Med. 2025;392:427–37. 10.1056/NEJMoa2410027. . PMID 39555826.
51. Knowler WC, Barrett-Connor E, Fowler SE, Hamman RF, Lachin JM, Walker EA, et al. Long-term effects and effect heterogeneity of lifestyle and metformin interventions on type 2 diabetes

incidence over 21 years in the US Diabetes Prevention Program randomised clinical trial. Lancet Diabetes Endocrinol. 2025;13:469–81. 10.1016/S2213-8587(25)00062-6. . PMID 40311647.

52. Newsome PN, Buchholtz K, Cusi K, Loomba R, Ratziu V, Sanyal AJ, et al. A Placebo-Controlled Trial of Subcutaneous Semaglutide in Nonalcoholic Steatohepatitis. N Engl J Med. 2021;384:1113–24. 10.1056/NEJMoa2028395. . PMID 33185364.
53. Kang Q, Xu J, Luo H, Tan N, Chen H, Cheng R, et al. Direct antiviral agent treatment leads to rapid and significant fibrosis regression after HCV eradication. J Viral Hepat. 2021;28:1284–92. 10.1111/jvh.13558. . PMID 34105867.
54. Holman RR, Paul SK, Bethel MA, Matthews DR, Neil HA. 10-year follow-up of intensive glucose control in type 2 diabetes. N Engl J Med. 2008;359:1577–89. 10.1056/NEJMoa0806470. . PMID 18784090.
55. Turashvili G, Gao Y, Ai DA, Ewaz AM, Gjeorgjievski SG, Wang Q, et al. Low interobserver agreement among subspecialised breast pathologists in evaluating HER2-low breast cancer. J Clin Pathol. 2024;77:815–21. 10.1136/jcp-2023-209055. . PMID 37714693.
56. Ai D, Turashvili G, Gjeorgjievski SG, Wang Q, Ewaz AM, Gao Y, et al. Subspecialized breast pathologists have suboptimal interobserver agreement in Ki-67 evaluation using 20% as the cutoff. Breast Cancer Res Treat. 2024;204:415–22. 10.1007/s10549-023-07197-3. . PMID 38157098.
57. Meira I, Menino J, Ferreira P, Leite AR, Gonçalves J, Ferreira HU, et al. Diabetes Remission After Bariatric Surgery: A 10-Year Follow-Up Study. Obes Surg. 2025;35:161–9. 10.1007/s11695-024-07592-9. . PMID 39715944.
58. Yuan M, Zhang Y, Liu B, Zhou L, Zhu Y, Li Z, et al. Treatment-free remission in chronic myeloid leukaemia patients with accelerated phase or tyrosine kinase inhibitor therapy failure. Br J Haematol. 2025;207:1408–15. 10.1111/bjh.70080. . PMID 40799021.
59. Courcoulas AP, Patti ME, Hu B, Arterburn DE, Simonson DC, Gourash WF et al. Long-Term Outcomes of Medical Management vs Bariatric Surgery in Type 2 Diabetes. JAMA. 2024;331:654–664. 10.1001/jama.2024.0318. PMID 38411644.
60. Nuti S, Zhang Y, Zerrouki N, Roach C, Bänfer G, Kumar GL, et al. High interobserver and intraobserver reproducibility among pathologists assessing PD-L1 CPS across multiple indications. Histopathology. 2022;81:732–41. 10.1111/his.14775. . PMID 35993150.
61. Gouda MA, Janku F, Yuan Y, Drusbosky LM, Chen AP, Zheng X, et al. Concordance between Tumor Tissue and Plasma DNA Genotyping in the NCI-MATCH Trial (EAY131). Clin Cancer Res. 2025;31:4299–310. 10.1158/1078-0432.CCR-24-3531. . PMID 40388547.
62. Yu C, Han Y, Wang M, Hua P, Zhang Y, Wang B. Concordance of ctDNA and tissue mutations in NSCLC: A meta-analysis. Cell Mol Biol (Noisy-le-grand). 2023;69:89–95. 10.14715/cmb/2023.69.8.14. PMID 37715416.
63. Jamal F, Sahwan O, Abidoye S, Alsulaiman AA, Sheashaa H, Ahn DH, et al. Concordance of Circulating Tumor DNA, Primary and Metastatic Tissue DNA, and Immunohistochemistry in

Gastroesophageal Carcinoma. JCO Precis Oncol. 2026;10:e2501194. 10.1200/PO-25-01194. . PMID 41881066.

64. Su H, Fan R, Tian M, Li Y, Wang Y, Wang T, et al. Pretreatment circulating tumor DNA and tissue genomic profiling concordance in ovarian cancer: a prospective observational study of influencing factors and clinical implications. Ther Adv Med Oncol. 2025;17:17588359251399468. 10.1177/17588359251399468. . PMID 41362586.
65. Seo SH, Park YS, Nam SK, Lee HS, Park DJ, Park KU. Concordance of circulating tumor DNA and matched formalin-fixed paraffin-embedded tumor tissue in gastric cancer as a predictor of recurrence. Korean J Clin Oncol. 2023;19:45–51. 10.14216/kjco.23009. . PMID 38229488.
66. French H, Mark Dalzell A, Srinivasan R, El-Matary W. Relapse rate following azathioprine withdrawal in maintaining remission for Crohn's disease: a meta-analysis. Dig Dis Sci. 2011;56:1929–36. 10.1007/s10620-011-1671-5. . PMID 21476031.
67. Kennedy NA, Kalla R, Warner B, Gambles CJ, Musy R, Reynolds S, et al. Thiopurine withdrawal during sustained clinical remission in inflammatory bowel disease: relapse and recapture rates, with predictive factors in 237 patients. Aliment Pharmacol Ther. 2014;40:1313–23. 10.1111/apt.12980. . PMID 25284134.
68. van Gerven NM, Verwer BJ, Witte BI, van Hoek B, Coenraad MJ, van Erpecum KJ et al. Relapse is almost universal after withdrawal of immunosuppressive medication in patients with autoimmune hepatitis in remission. J Hepatol. 2013;58:141-7. 10.1016/j.jhep.2012.09.009. PMID 22989569.
69. Zhao M, Zhang P, Fang L, Luo Z, Gu J, Lin Z. Possible predictors for relapse from etanercept discontinuation in ankylosing spondylitis patients in remission: a three years' following-up study. Clin Rheumatol. 2018;37:87–92. 10.1007/s10067-017-3763-x. . PMID 28785857.
70. Lee SE, Choi SY, Song HY, Kim SH, Choi MY, Park JS, et al. Imatinib withdrawal syndrome and longer duration of imatinib have a close association with a lower molecular relapse after treatment discontinuation: the KID study. Haematologica. 2016;101:717–23. 10.3324/haematol.2015.139899. . PMID 26888022.
71. Zaakouk M, Quinn C, Provenzano E, Boyd C, Callagy G, Elsheikh S et al. Concordance of HER2-low scoring in breast carcinoma among expert pathologists in the United Kingdom and the republic of Ireland -on behalf of the UK national coordinating committee for breast pathology. Breast. 2023;70:82–91. 10.1016/j.breast.2023.06.005. PMID 37419078.
72. Mills AM, Bennett JA, Banet N, Watkins JC, Kundu D, Pinto A. Interobserver Agreement on the Interpretation of Programmed Death-ligand 1 (PD-L1) Combined Positive Score (CPS) Among Gynecologic Pathologists. Am J Surg Pathol. 2023;47:889–96. 10.1097/PAS.0000000000002058. . PMID 37272261.
73. Sy A, Alkashash A, Drage MG, Machacek ME, Shih AR, Shroff SG, et al. Claudin 18.2 immunohistochemistry in gastroesophageal adenocarcinomas: A study of interobserver variability among gastrointestinal pathologists. Hum Pathol. 2026. 10.1016/j.humpath.2026.106279. . PMID 42822733.

74. Kyuno D, Yanazume K, Saito AC, Ono Y, Ito T, Imamura M, et al. Concordance of claudin-18.2 expression in biopsy, resection, and recurrent specimens: implications for zolbetuximab therapy in pancreatic ductal adenocarcinoma. Tissue Barriers. 2026;14:2535047. 10.1080/21688370.2025.2535047. . PMID 40694660.
75. Emsley R, Nuamah I, Gopal S, Hough D, Fleischhacker WW. Relapse After Antipsychotic Discontinuation in Schizophrenia as a Withdrawal Phenomenon vs Illness Recurrence: A Post Hoc Analysis of a Randomized Placebo-Controlled Study. J Clin Psychiatry. 2018;79. 10.4088/JCP.17m11874. . PMID 29924507.
76. Wei Q. Lung cancer origin: a seven-state linear model: pulmonary power target combinations, SEER calibration, and state-resolved readouts. Res Square. 2026. 10.21203/rs.3.rs-11655637/v1.
77. Wei Q. Liver cancer origin: a seven-state linear model: OST triad power target combinations, SEER calibration, and state-resolved readouts. Res Square. 2026. 10.21203/rs.3.rs-11655621/v1.
78. Wei Q. Breast cancer origin: a nine-state linear model: B3 two-axis power target combinations, SEER calibration, and state-resolved readouts. Res Square. 2026. 10.21203/rs.3.rs-11655400/v1.
79. Wei Q. Ovarian cancer origin: a seven-state linear model: five-group power target combinations, SEER calibration, and state-resolved readouts. Res Square. 2026. 10.21203/rs.3.rs-11655513/v1.
80. Wei Q, Ma YT, Zhu GJ. The origins of childhood cancer. Res Square. 2026. 10.21203/rs.3.rs-11626154/v1.
81. Wei Q, Ma YT, Zhu GJ. Three permissiveness gates multiply to determine mutation realisation: age-specific incidence and the dialysis-transplant separation from one parameter set. Res Square. 2026. 10.21203/rs.3.rs-11544766/v1.
82. Wei Q, Ma YT, Zhu GJ. Why childhood cancer arises without a precancerous sequence: evidence, gaps, and a measurement agenda. Preprints org. 2026. 10.20944/preprints202610.2429.v1.

Figures

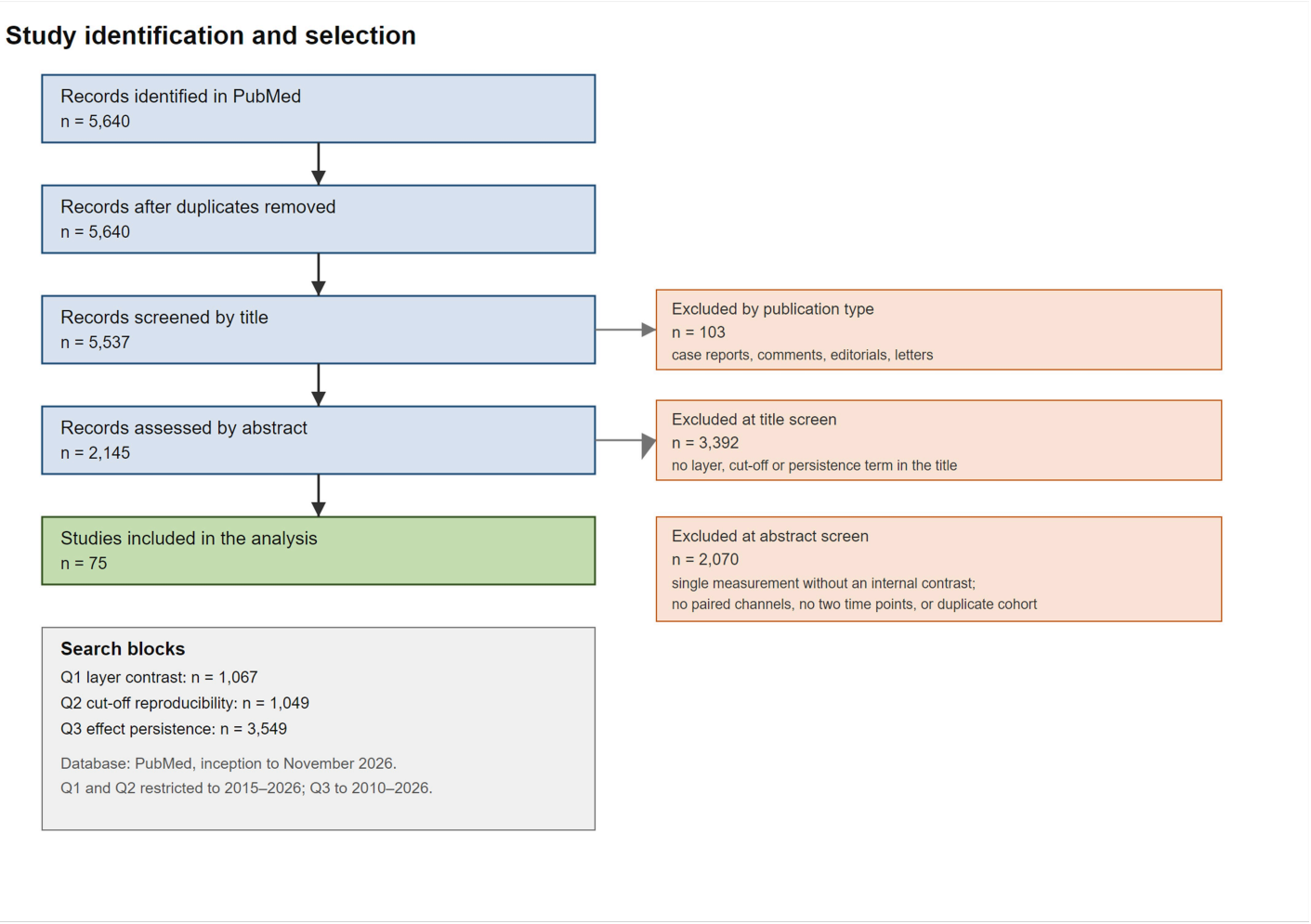

Figure 1

Identification and selection of studies.

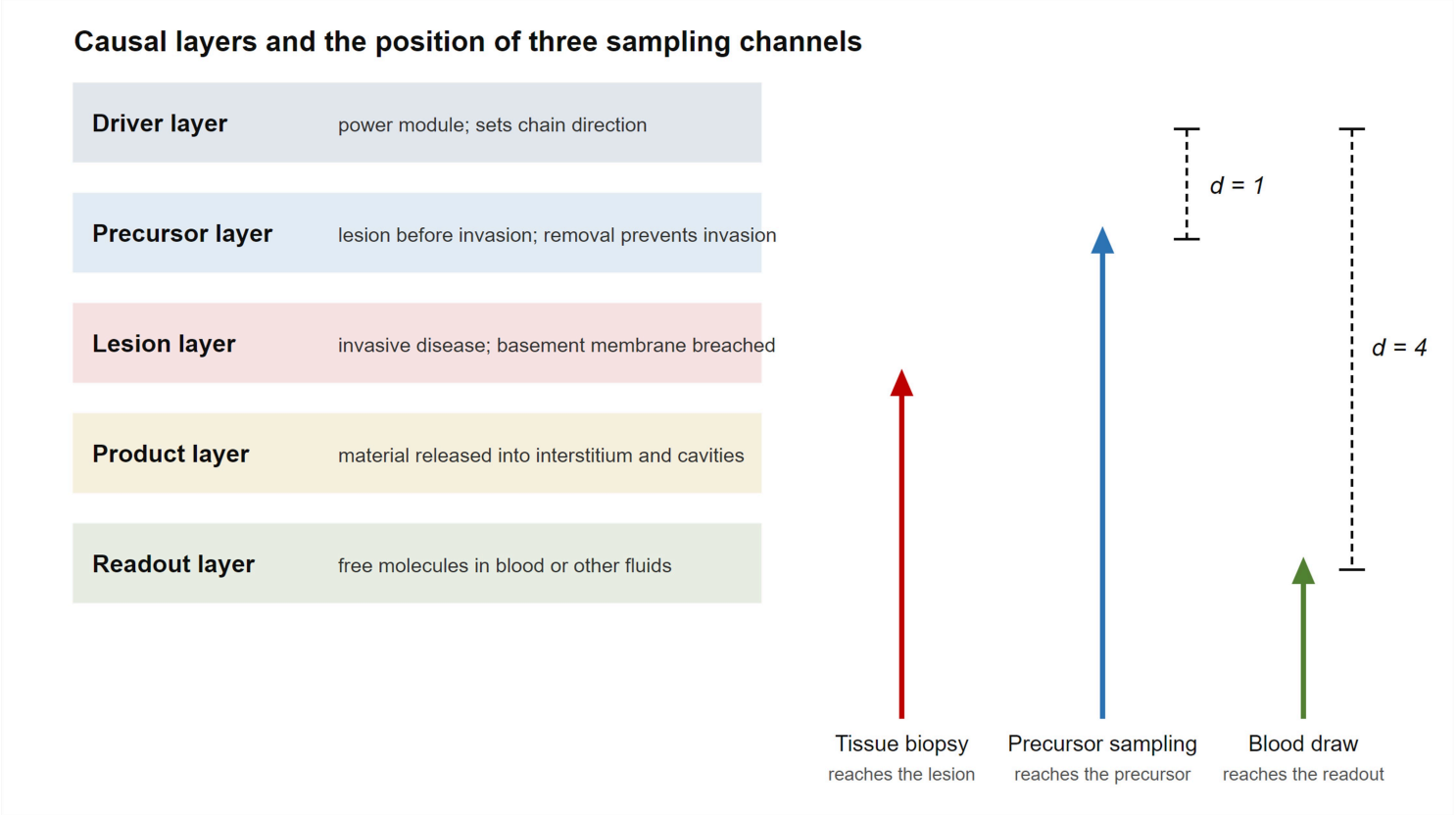

Figure 2

Layers of the causal chain and three sampling channels.

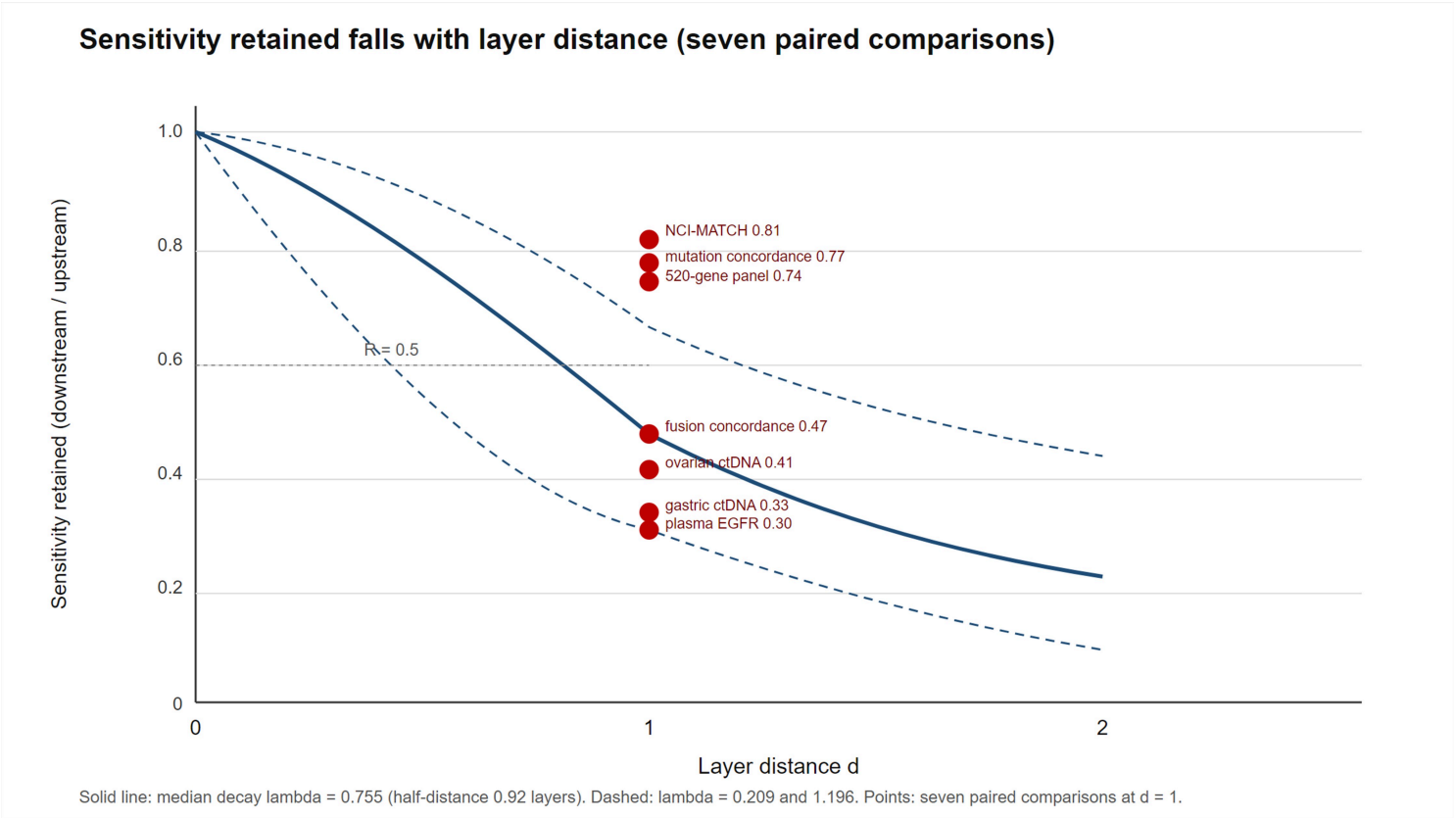

Figure 3

Sensitivity retained against layer distance.

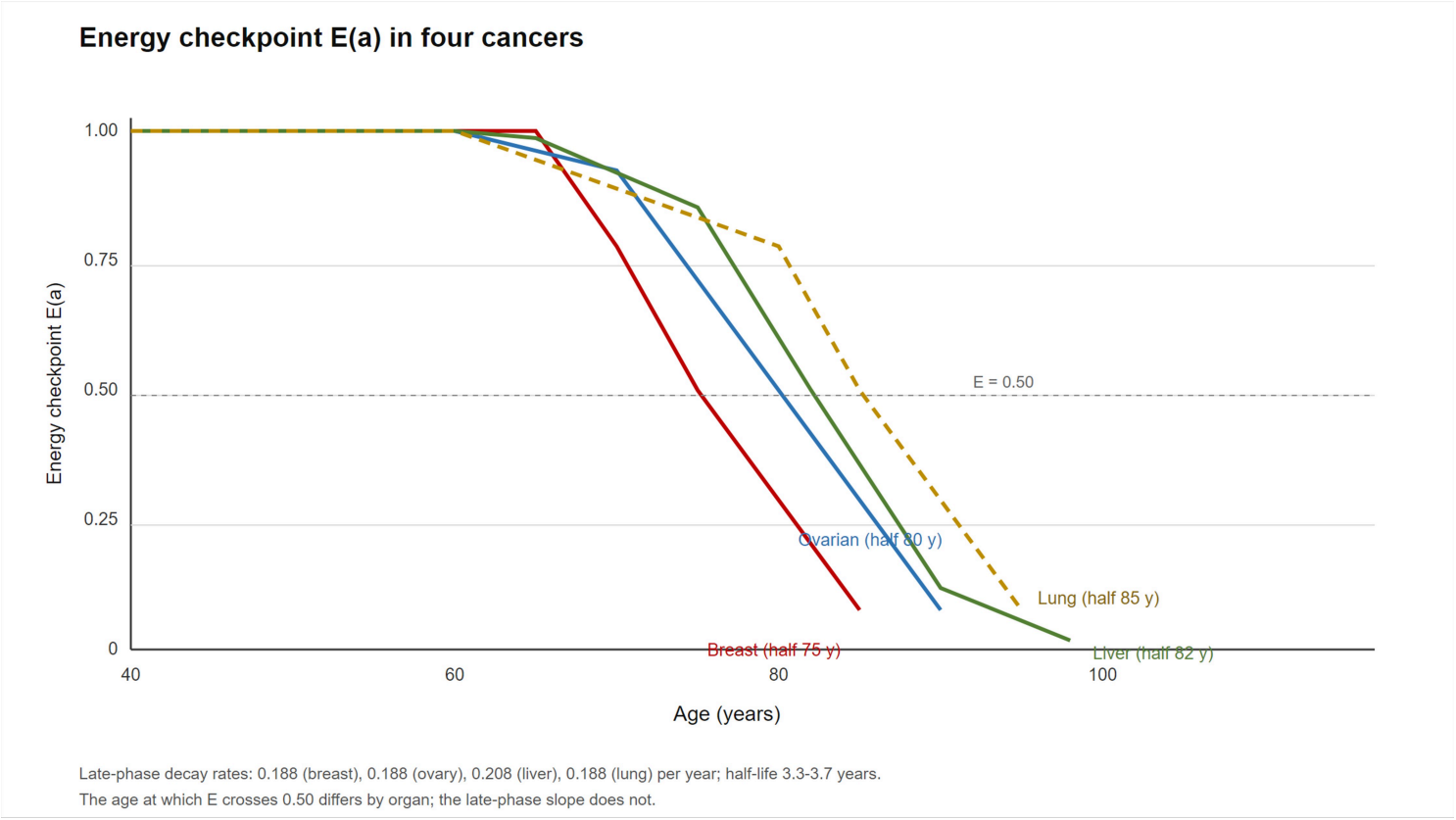

Figure 4

Energy checkpoint E(a) in four cancers.

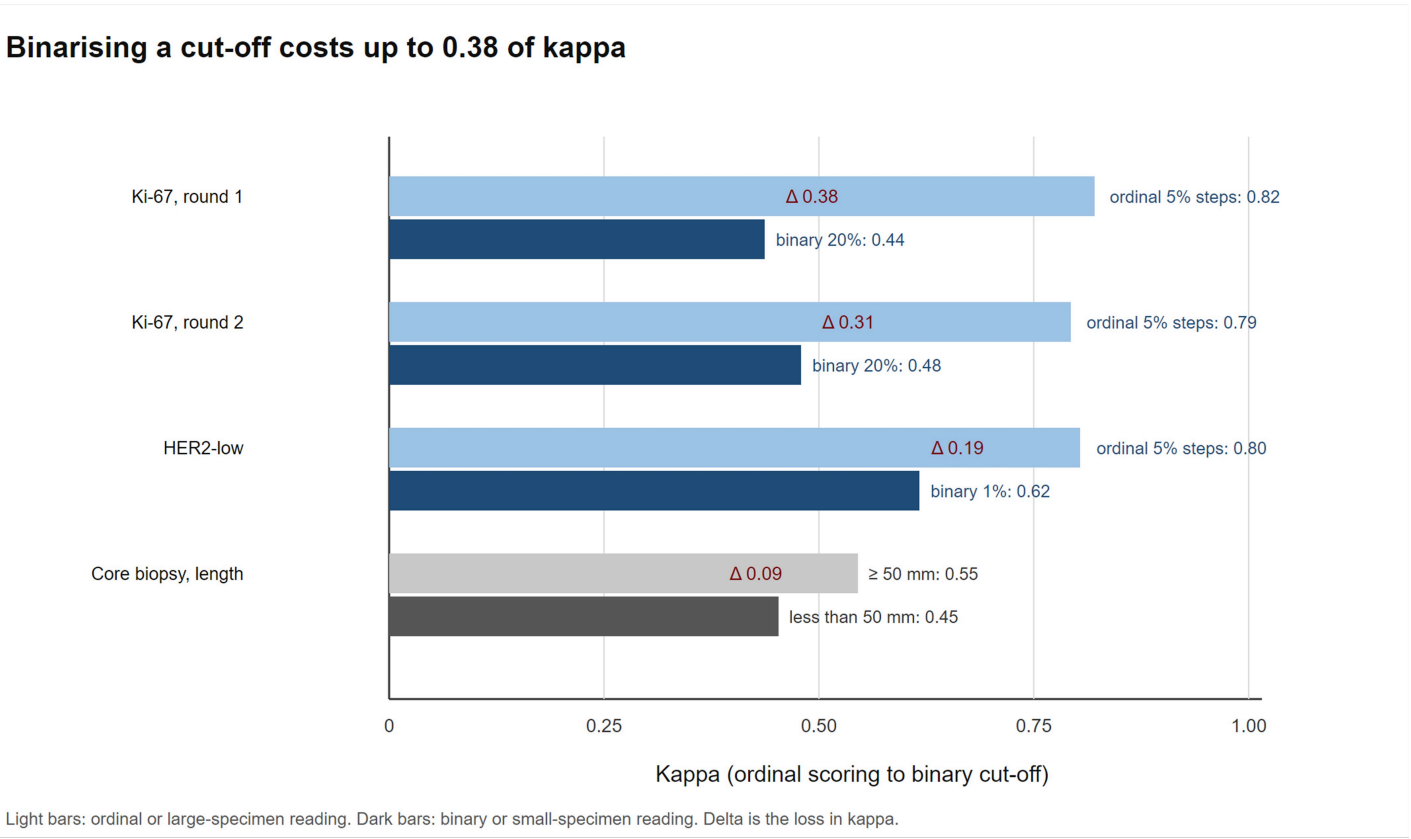

Figure 5

Binarisation and specimen-size losses in kappa.

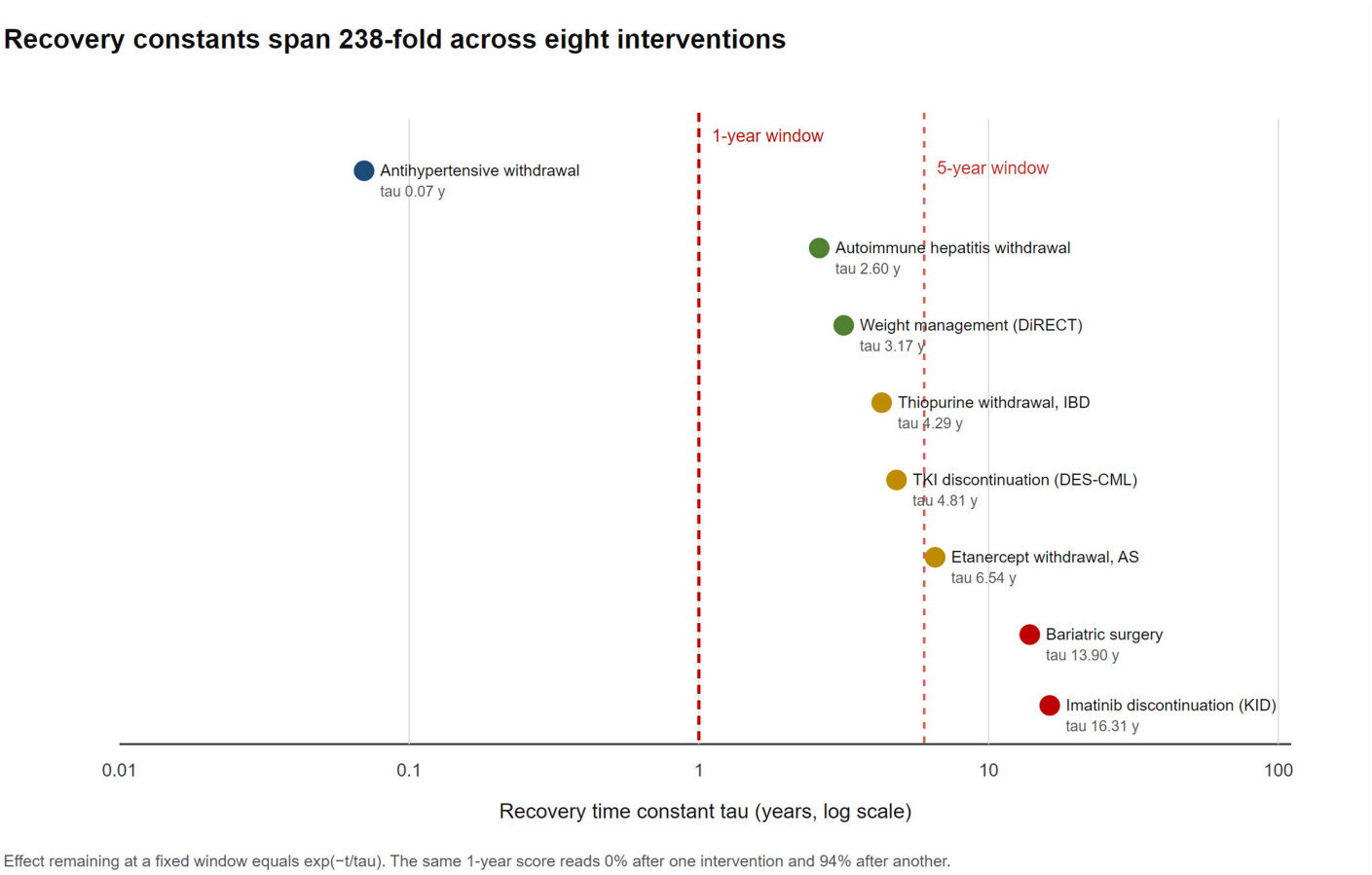

Figure 6

Recovery time constants on a logarithmic scale with one-year and five-year windows.