

Modelling for Personalised Screening and Prevention: International Experience

Universities

Clinicians

Professor Amy Ming-Fang Yen

Taipei Medical University

7th Oct 2026

NGO

Community

Enterprise



ROADMAP

The Arc of This Talk

Two questions: what is hidden in screening data, and how do we act on it for each woman?

01

Hidden States & Digital Twin



02

Precision Screening

1. Hidden States & Digital Twin

Screening sees only a sliver of the disease process.

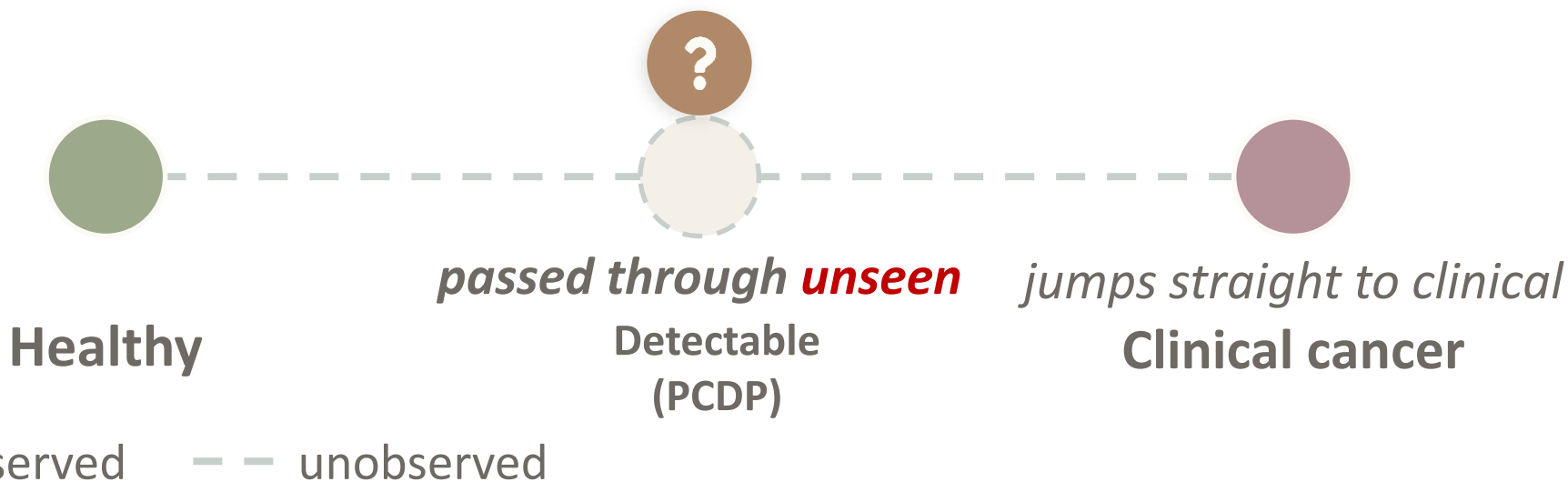
Stochastic models recover the hidden transitions, and a digital twin rebuilds the complete trajectory of every woman.

What Screening Data Cannot see?

1 Screen-detected, then removed



2 Interval cancer — surfaces between screens



What Screening Data Cannot see?

1 Screen-detected, then removed

removed ✂

Two important factors in screening:
disease status (State) and time.

we never see the rest of its natural history

2 Interval between successive screens

→ **Stochastic Process**

?

*passed through **unseen***

jumps straight to clinical

Healthy

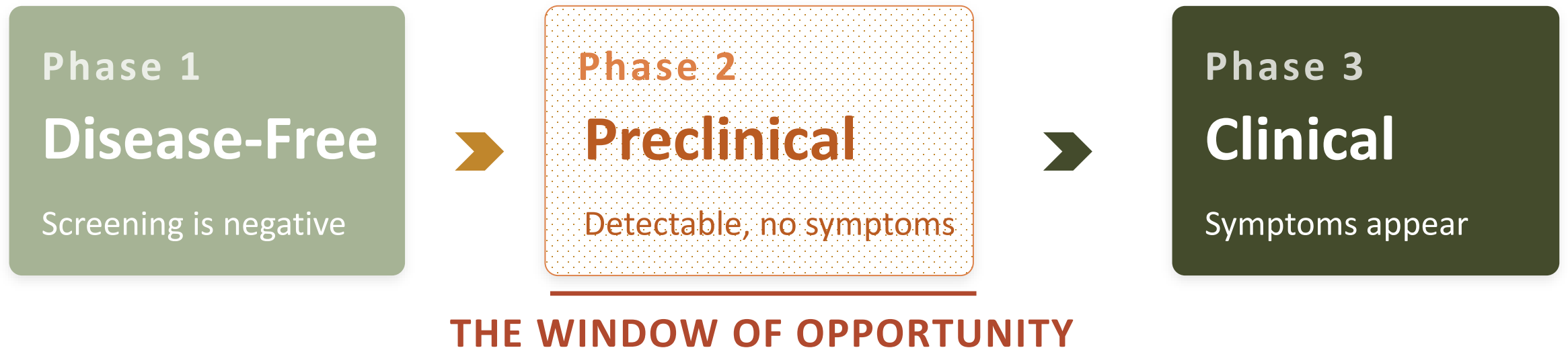
Detectable
(PCDP)

Clinical cancer

— observed - - - unobserved

WINDOW OF OPPORTUNITY

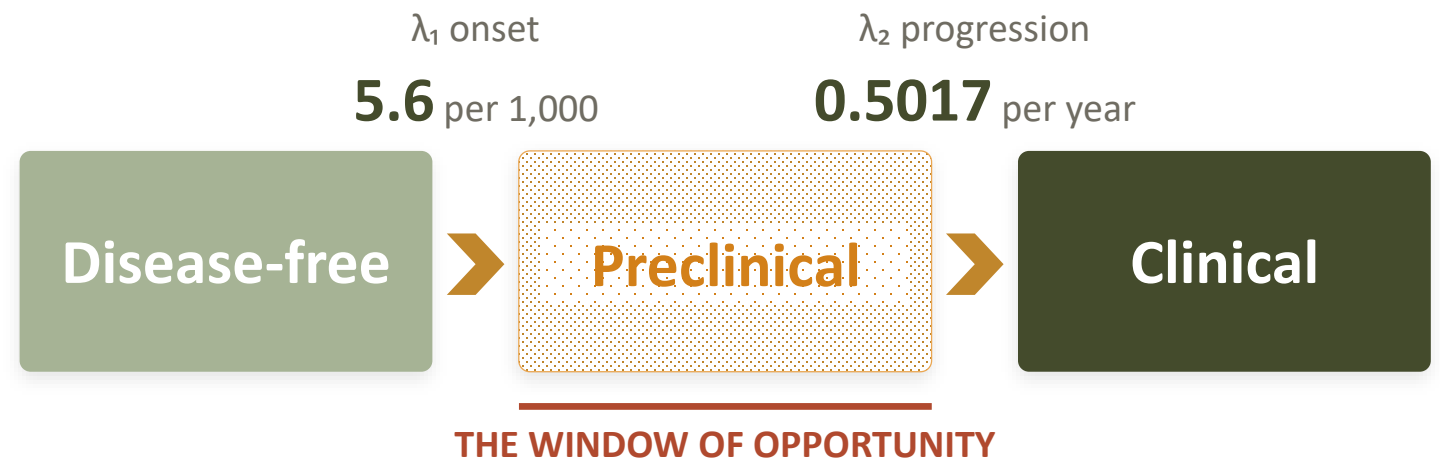
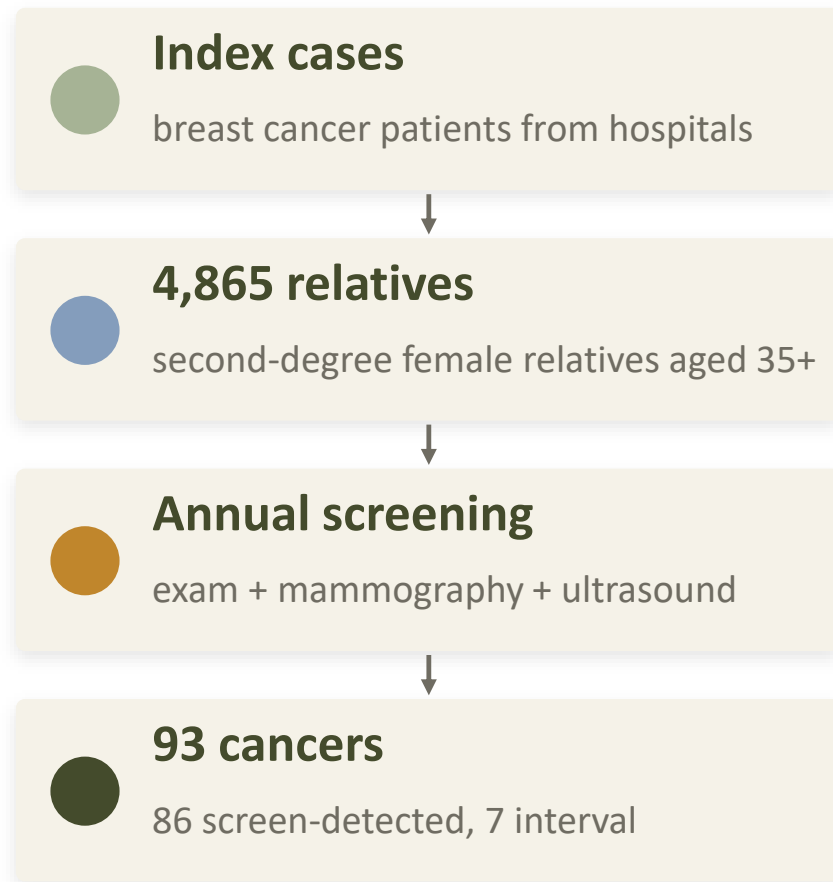
Where Screening Can Act



***Sojourn time** = how long disease stays detectable but silent in Phase 2*

Estimating the Hidden Window in High-Risk Women

Applying the stochastic process to family-based breast cancer screening in Taiwan, 1995–1998



1.99

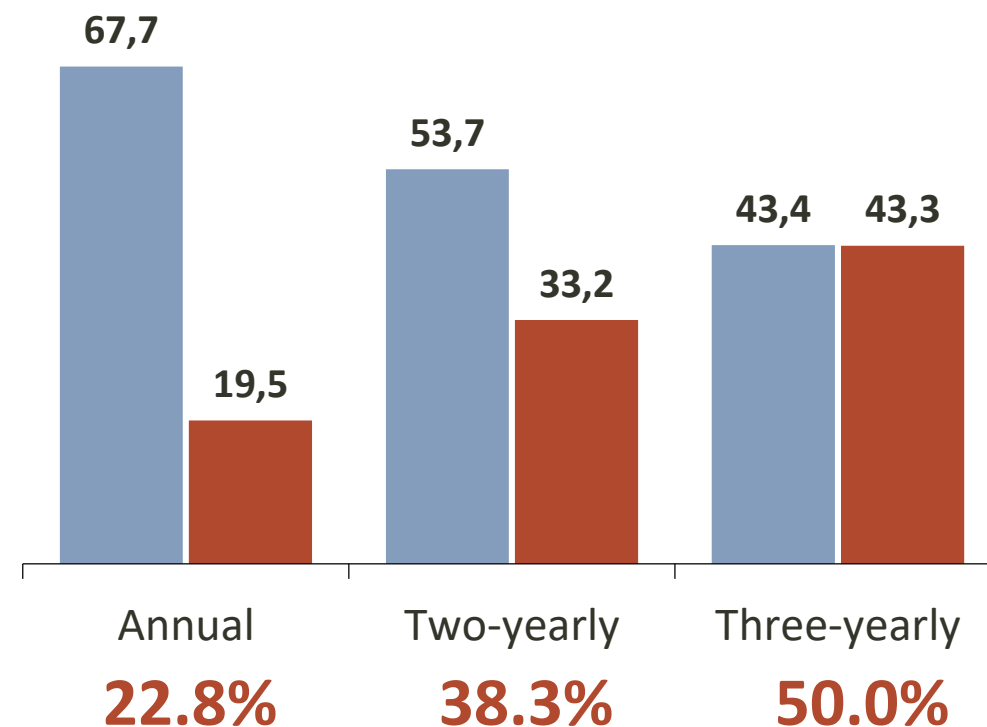
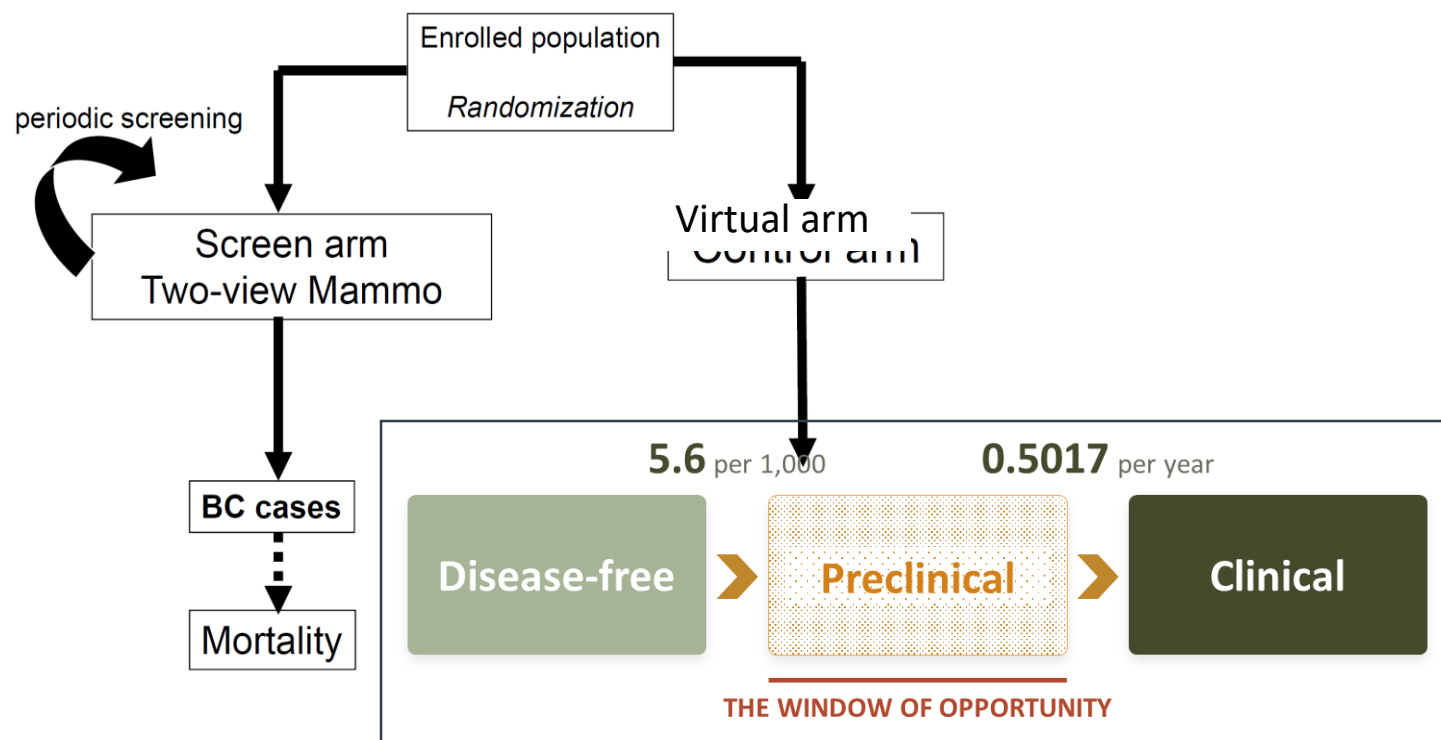
years mean sojourn time

APPLICATION · STOCHASTIC TWIN GENERATOR

Empirical data:
Taiwan

Simulating Screening Regimes

TAMCAS parameters drive virtual twins screened annually, every two or every three years



Interval cancers as a share of incident cancers, (B)/(A+B)

■ Incident screen-detected (A) ■ Interval cancers (B)

APPLYING THE DIGITAL TWIN TO AN RCT

Empirical data:
Sweden

Deciphering the Swedish Two-County Trial

1977–1985

Randomised trial

Invited vs uninvited, with
perfect comparator

*W (Kopparberg/Dalarna) and
E (Östergötland) counties*

1986–2005

Service screening I

National screening policy

W (Kopparberg/Dalarna)

2006–2022

Service screening II

Adds imaging biomarkers,
ER, PR, HER2, basal

W (Kopparberg/Dalarna)



78,085 (ASP)

invited women
active study population

56,782 (PSP)

uninvited women
passive study population

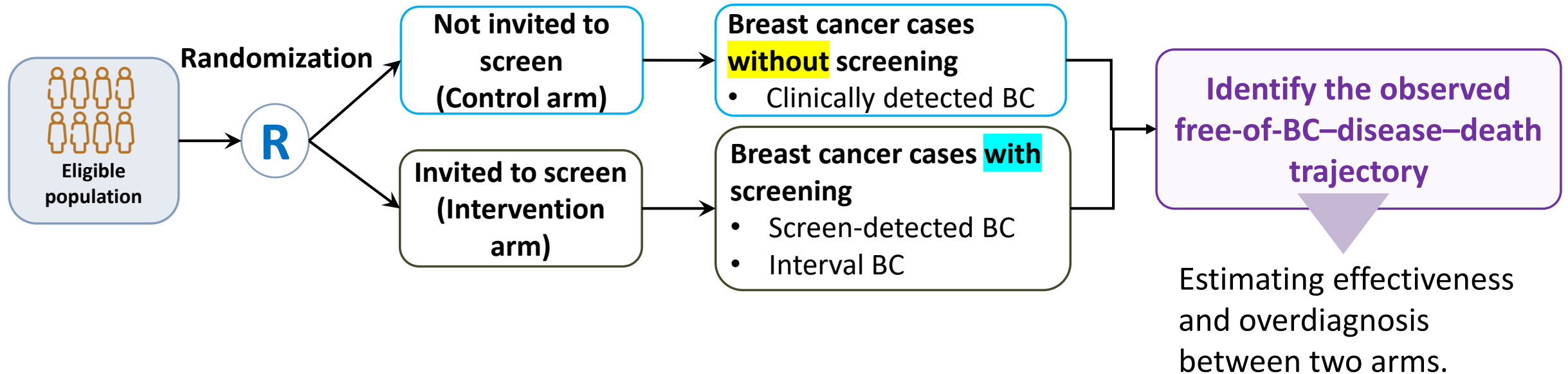
2,468

breast cancers
labelled by detection mode

THE CONVENTIONAL VIEW

A Randomised Trial Compares Arms at Population-level

Conventional RCT Design

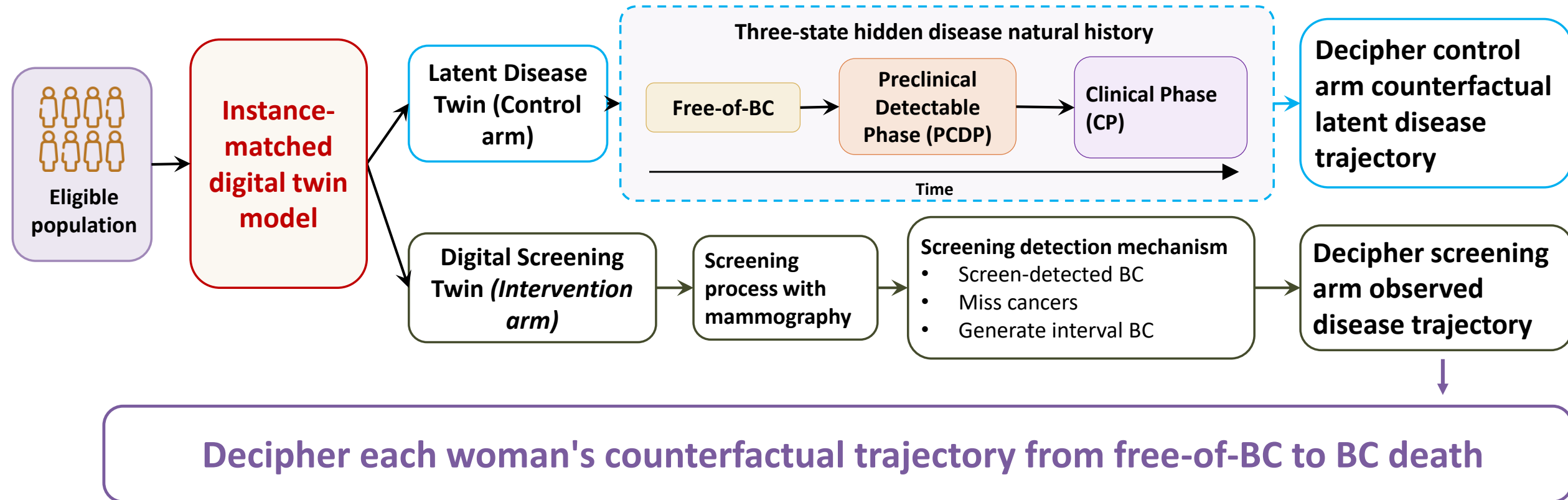


Randomisation guarantees an unbiased contrast, but averages over every woman and every biological pathway.

INSTANCE-MATCHED DIGITAL TWIN

Two Twins Decipher Biology and Detection

Instance-matched digital twin design

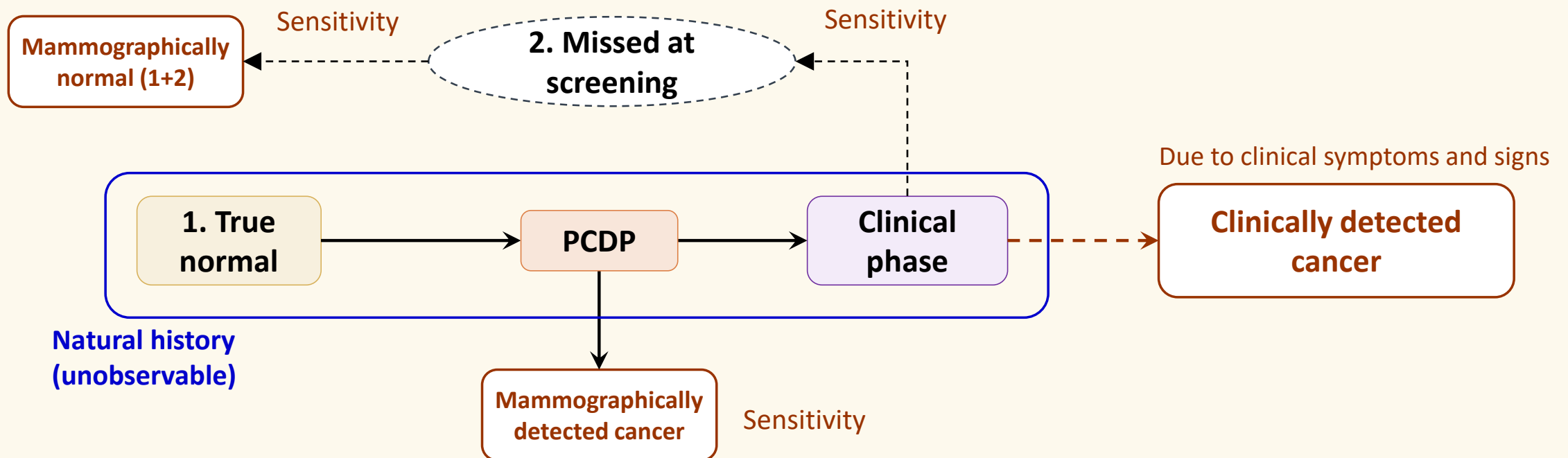


A trial answers a question about arms. The twin answers a question about people.

THE BIOLOGICAL ENGINE

Hidden History Behind the Observed Data

Three-state hidden disease natural history model and observed screening data



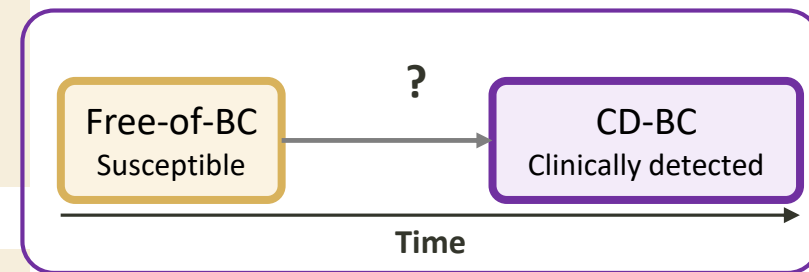
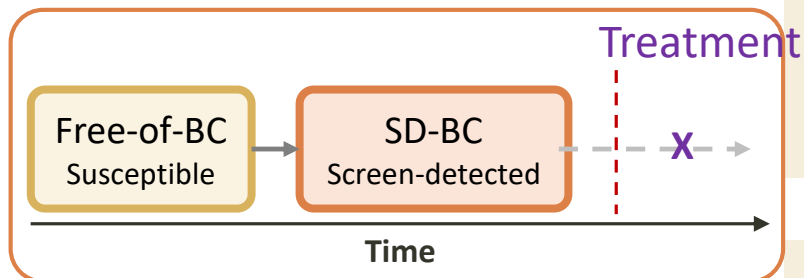
METaverse Synchronisation Engine

From a Partial Record to a Complete Trajectory

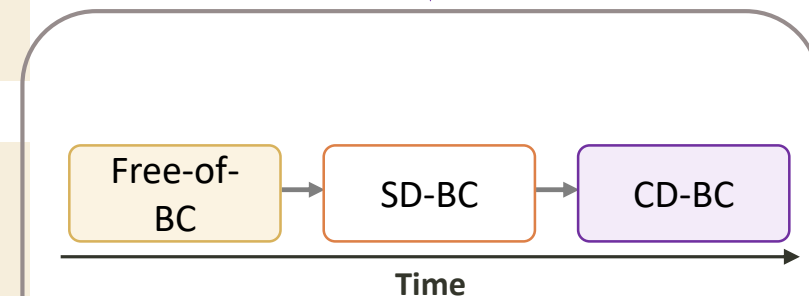
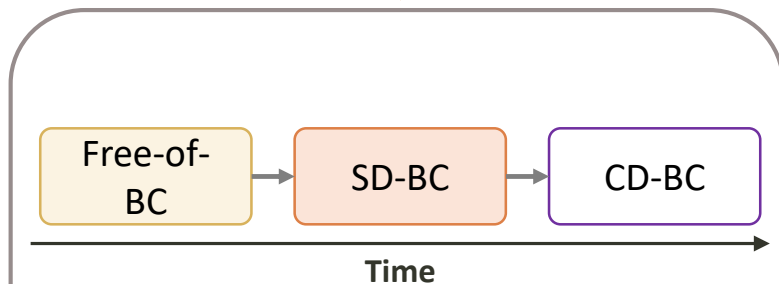
(A) Screen-detected breast cancer

(B) Interval breast cancer

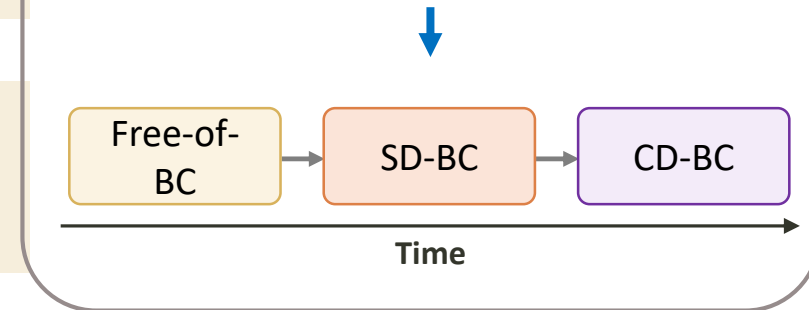
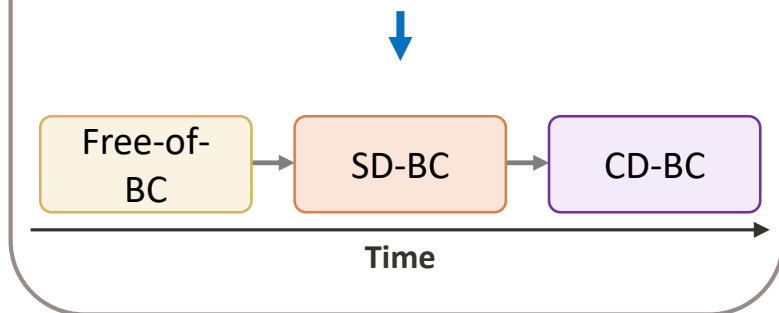
Observed
(real world)



Instance-matched
digital twin
(counterfactual
progression)



Reconstructed
complete
trajectory



1. Identify instance-
matched digital twin

2. Synchronize physical
and virtual trajectories

3. Infer latent disease
trajectory

4. Integrate trajectories

METaverse-SYNTHESIZED DISEASE THREAD

One Complete Natural History for Every Woman

Metaverse-synthesized digital disease thread

Integrating physical and virtual data into one complete natural history

Free-of-BC
(Susceptible)



Preclinical phase
(SD-BC)



**Counterfactual
progression**
(untreated course)



Clinical presentation
(CD-BC)

Complete natural history: Susceptible → PCDP → CP



Enables counterfactual evaluation and personalized screening optimization

Sojourn time and lead-time distribution

Programme sensitivity with I/E ratio

False positives and overdiagnosis

Long-term effectiveness: advanced cancer, mortality

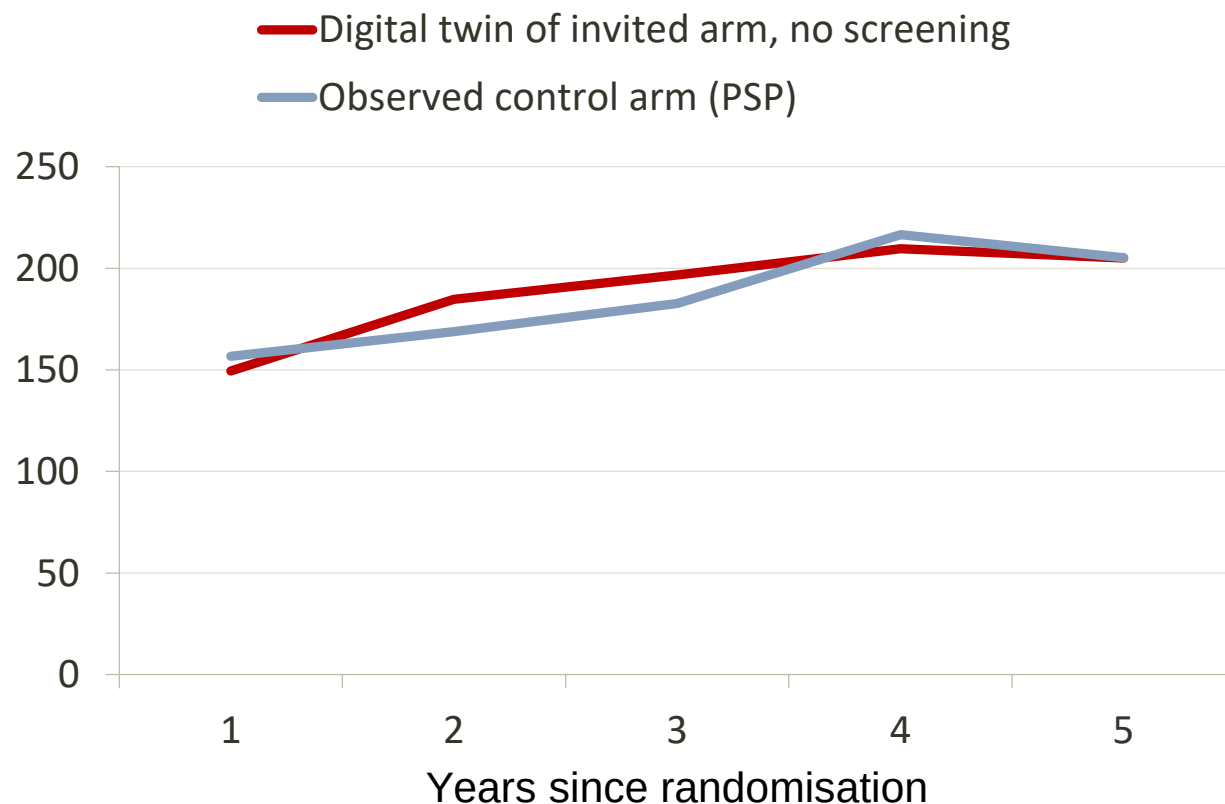
VALIDATION WITHIN THE TRIAL

Empirical data:
Swedish Two-County Trial

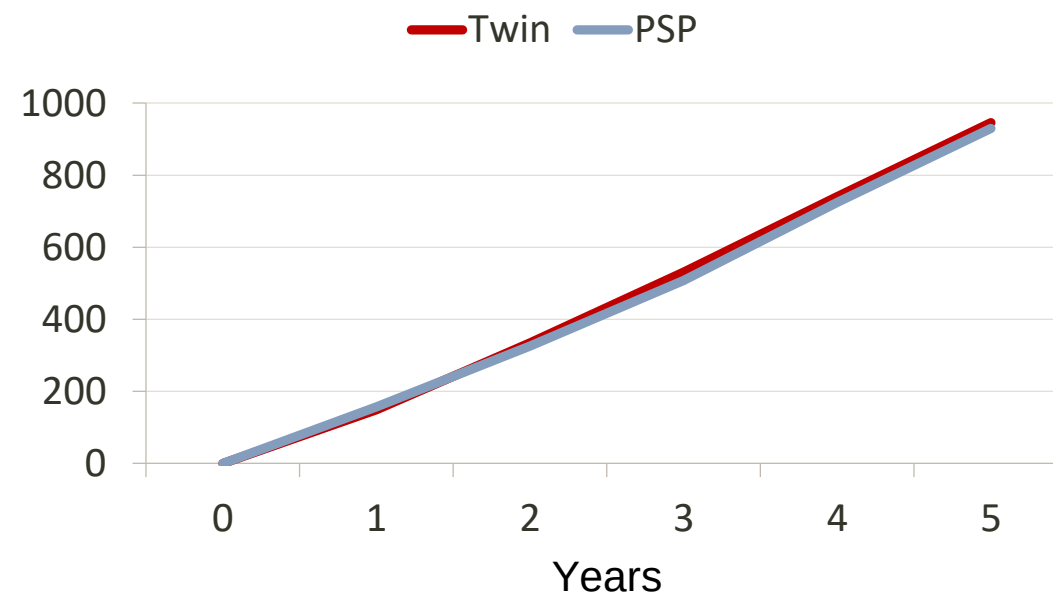
METaverse

The Twin Reproduces a Control Arm It Never Saw

Annual incidence per 100,000



Cumulative incidence per 100,000



945.7 vs 929.8 at 5 years

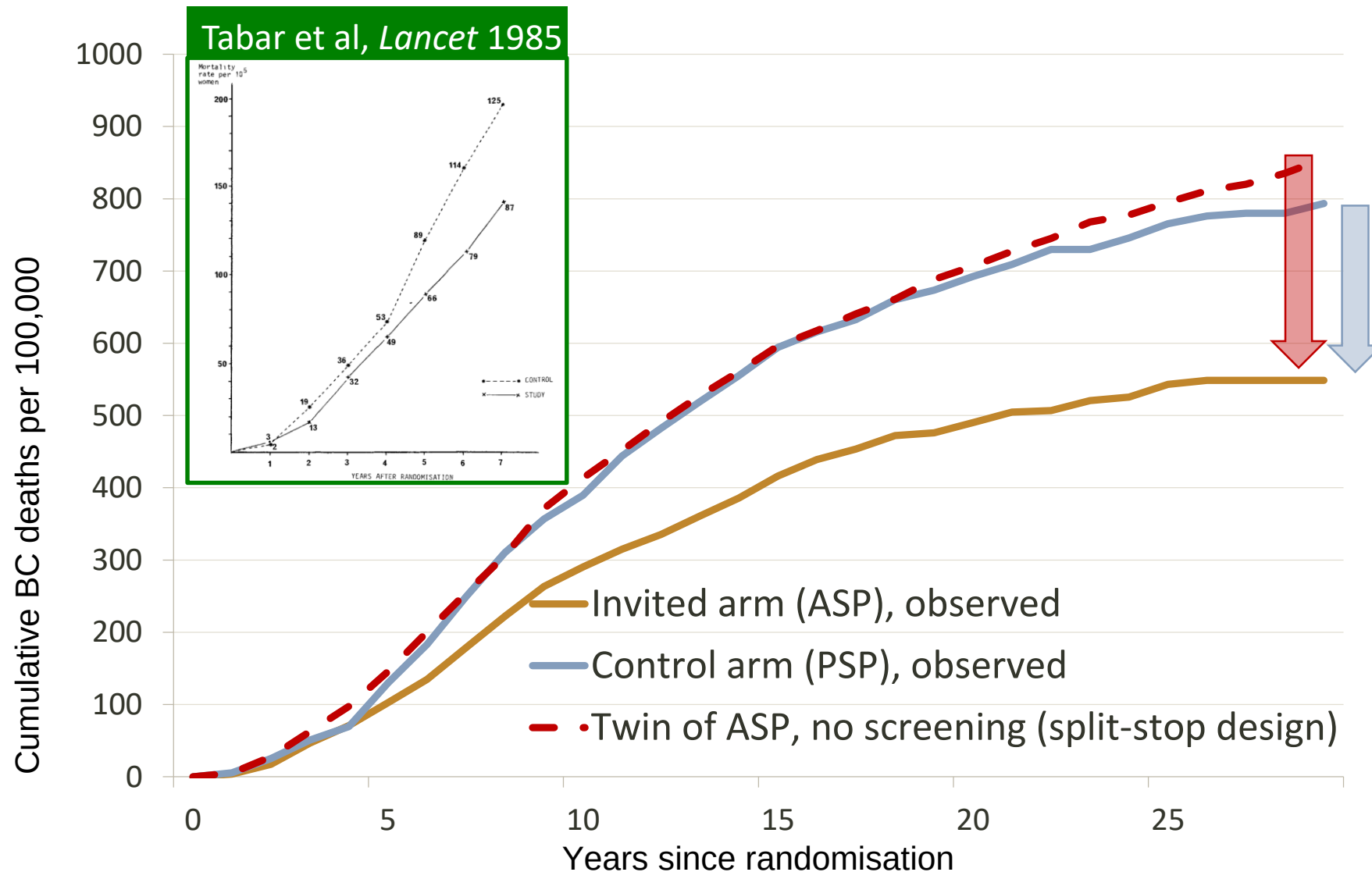
PSP inside the 95% replicate envelope every year

VALIDATION OF LONG-TERM MORTALITY

Empirical data:
Swedish Two-County Trial

The Twin Also Reproduces 29-Year Mortality

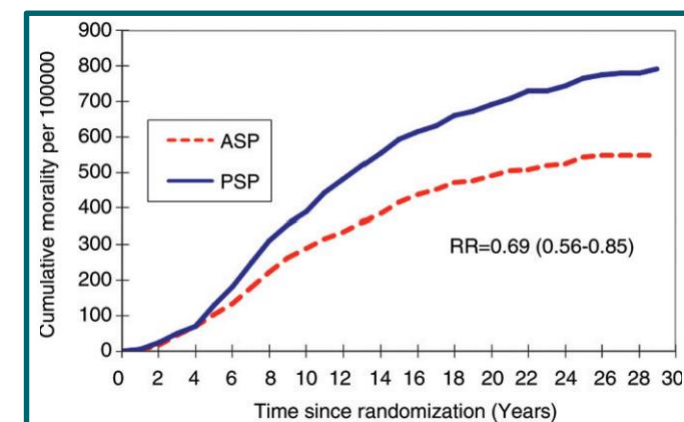
METAVERGE



Mortality reduction

33% (23–42) twin

31% (15–44) trial



Tabar et al, *Radiology* 2011

VALIDATION ACROSS TRIALS · MASAI

Same Biology, Different Sensitivity

Empirical data:
MASAI Trial, Sweden

Shared natural history

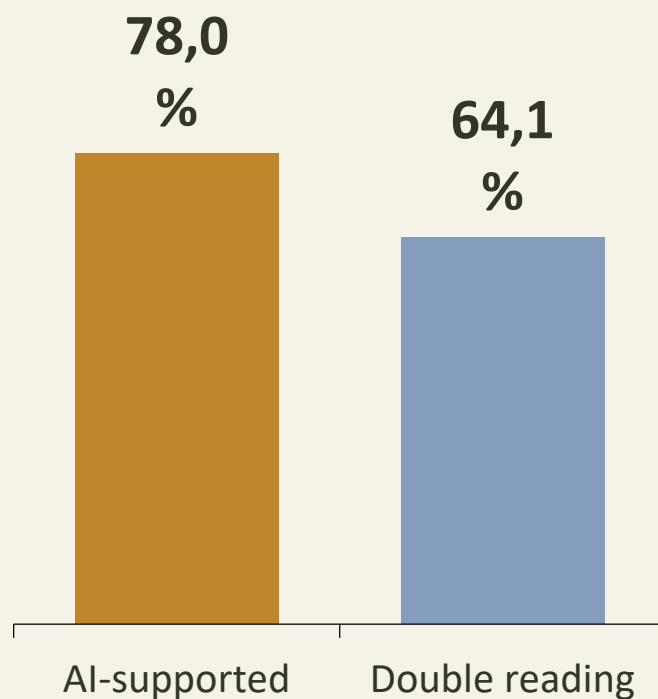
1.91

preclinical incidence per
1,000 person-years

4.16

years mean sojourn time

Test sensitivity



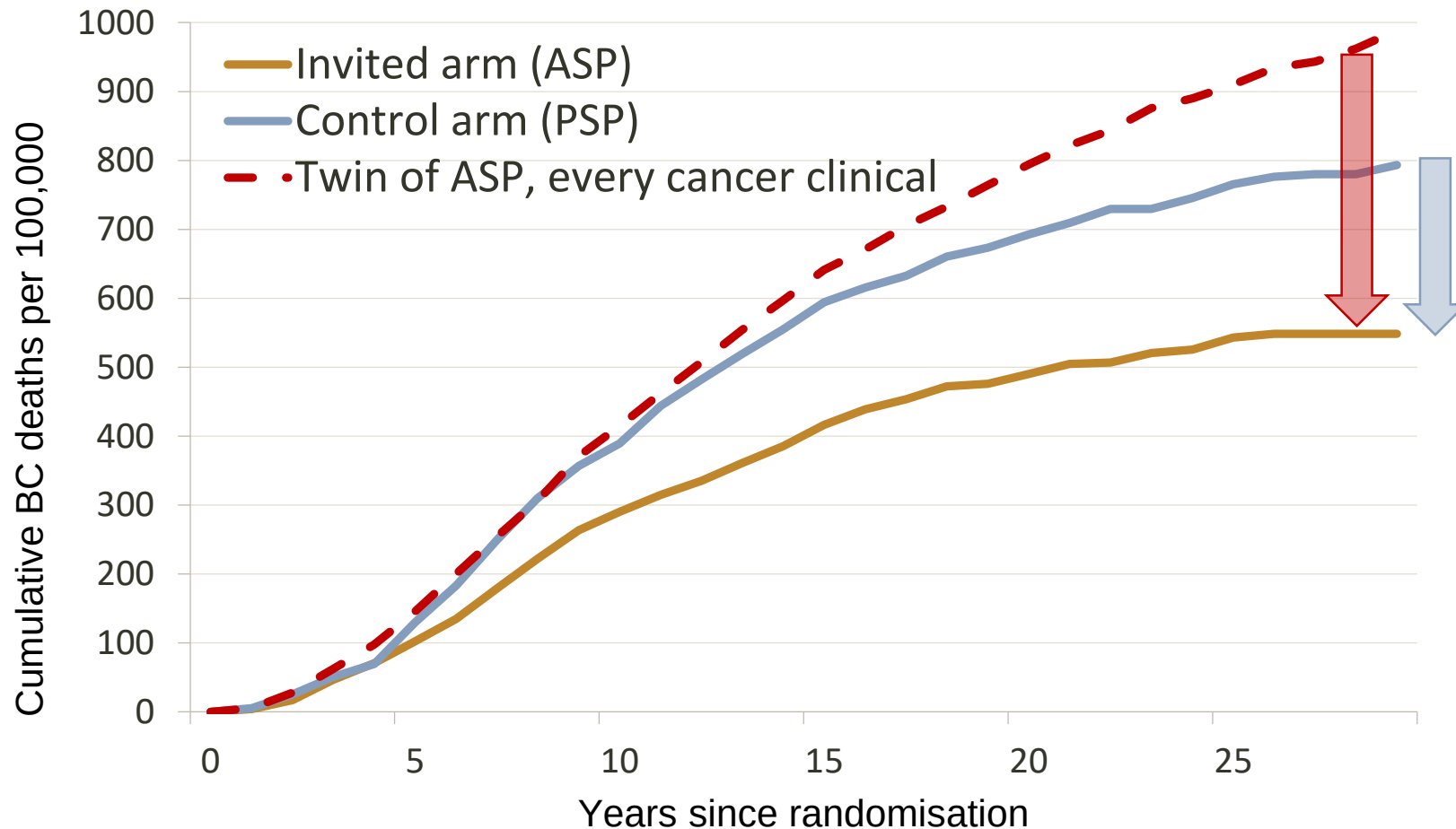
Projected vs observed

	model	trial
Screen-detected	599.2	600
Interval cancers	175.8	175
IC proportion, AI	18.9%	19.5%
IC proportion, DR	26.8%	26.2%

AI cuts interval cancers by seeing more of the same tumours, not by finding different ones.

LONG-TERM EFFECTIVENESS

Against a Truly Unscreened Twin, Mortality Falls 45%



Empirical data:
Swedish Two-County Trial

METAVVERSE

45%

548.5 vs 988.9: against its
own no-screening twin
(30% at 10 y, 38% at 20 y)

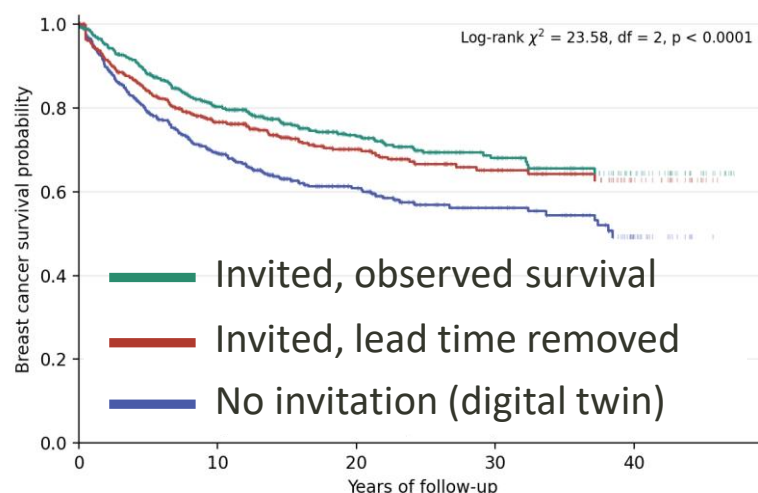
31%

against the control arm
(793.3), which received a
compensatory screen

Screening Cuts Breast Cancer Death

Invited women diagnosed during the trial, each compared with her own no-screening digital twin

All breast cancer

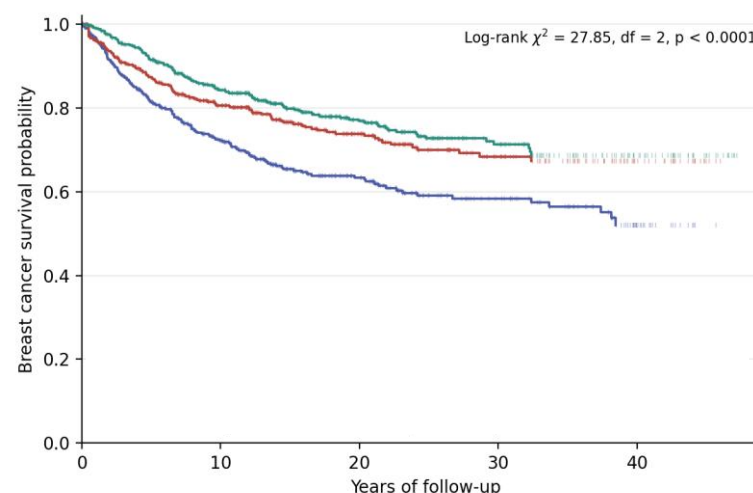


HR **0.714**

(0.579–0.881) lead time calibrated

observed **0.608**

Exposed to screening

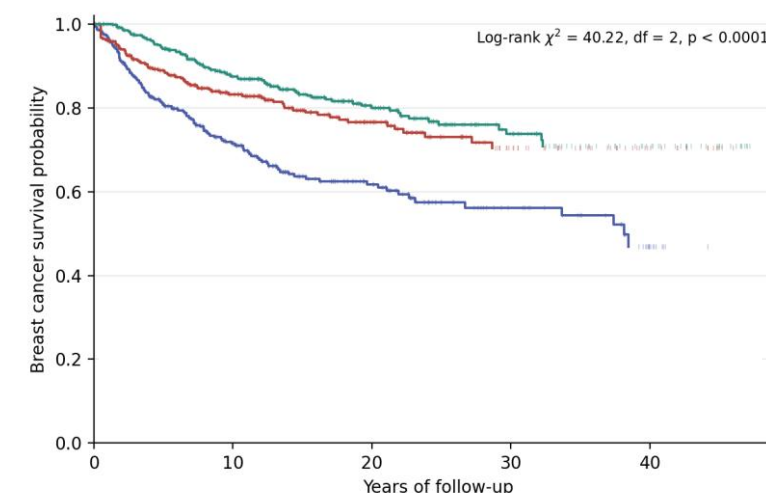


HR **0.658**

(0.520–0.833) lead time calibrated

observed **0.550**

Screen-detected



HR **0.542**

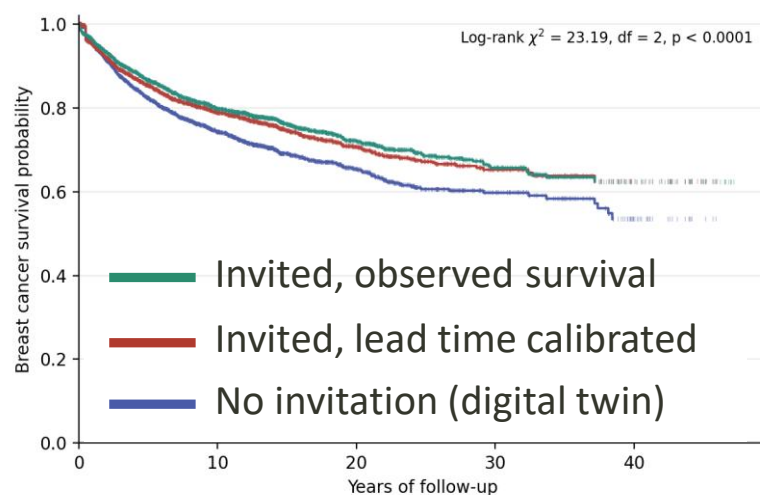
(0.407–0.723) lead time calibrated

observed **0.426**

The Benefit Holds With Service Screening to 2022 **METaverse**

Invited women diagnosed through 2022, each compared with her own no-screening digital twin

All breast cancer

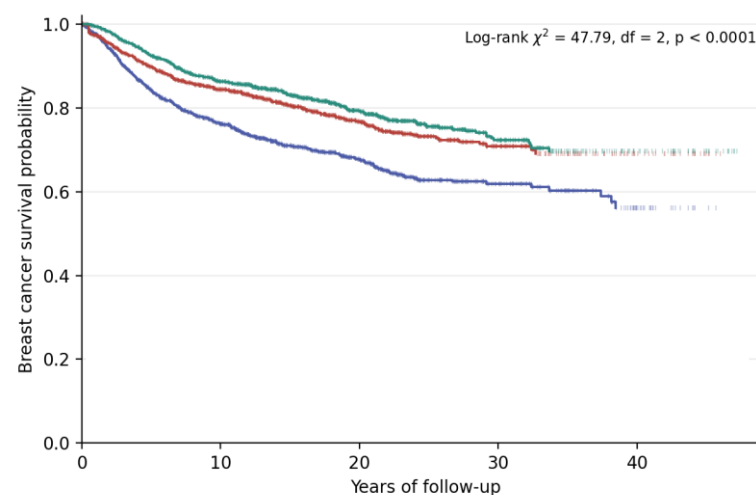


HR **0.812**

(0.721–0.915) lead time calibrated

observed **0.760**

Exposed to screening

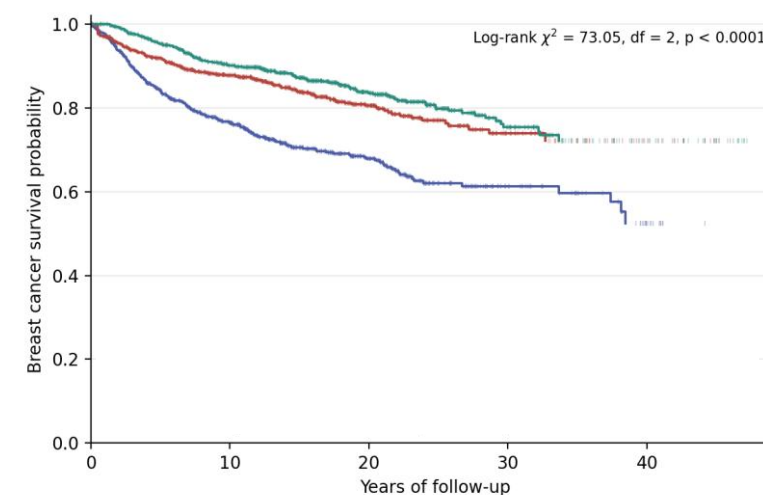


HR **0.663**

(0.560–0.785) lead time calibrated

observed **0.575**

Screen-detected



HR **0.535**

(0.433–0.662) lead time calibrated

observed **0.432**



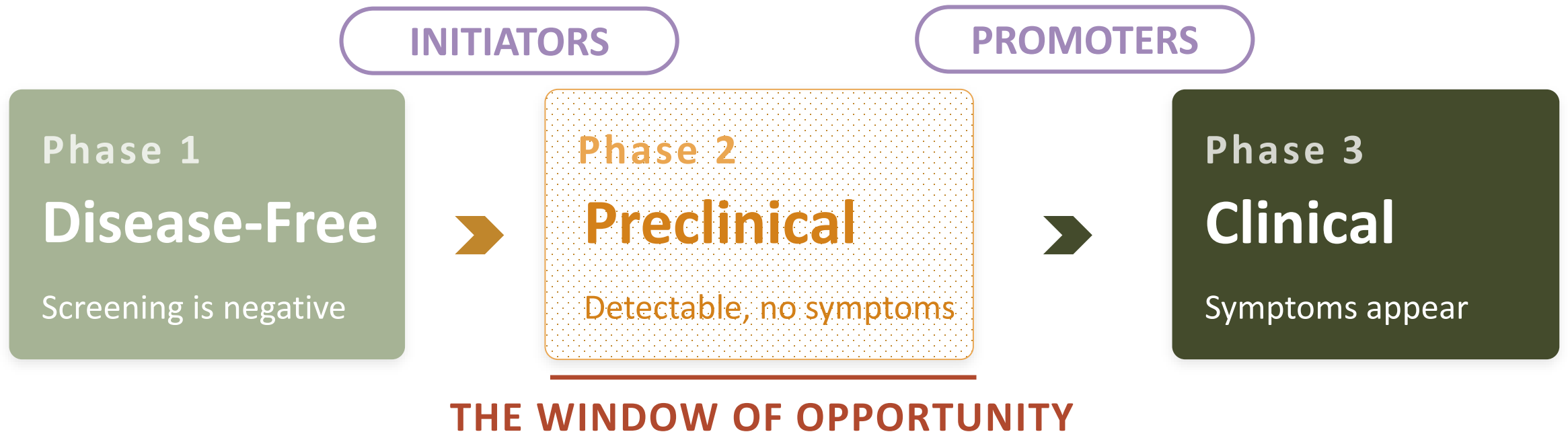
2. Precision Screening

Non-homogeneous rates, initiators and promoters, and tumour biology turn the model into an individualized screening strategy.



INITIATORS AND PROMOTERS

Towards Personalized Screening



Sojourn time = how long disease stays detectable but silent in Phase 2

COHORT ARCHITECTURE

50,666 Women, Two Eras

Yen AMF, Wu WYY, Tabar L, et al. J Epidemiol 2017;27:98–106

Empirical data:
Dalarna, Sweden

Era

Breast cancers

Covariates available

1977–1992
The Trial Period

Women aged 40–69 in
Kopparberg; screened every 24
months

1,321 cancers

789 SD

532 IN

SD screen-detected IN interval

Risk factors, trial participants

- Body mass index
- Breast density
- Age at menarche
- Age at first full-term pregnancy
- Menopausal status
- Family history (female relatives)

1993–2010
Service Screening

Registry linkage adds tumour
phenotypes and molecular
subtypes

1,614 cancers

1,135 SD

479 IN

SD screen-detected IN interval

IHC biomarkers, by year collected

ER	from 1996
PR	from 1996
HER2	1996–98; from 2002
Ki-67	1996–98; from 2006
Basal/myoepithelial	1996–98; from 2006

ADJUSTED RELATIVE RISKS

What Drives Each Force

Empirical data:
Dalarna, Sweden

Triggers of the preclinical phase

Family history **+89%**

Dense breast **+41%**

Late pregnancy **+23%**

Obesity **+15%**

Accelerators to clinical cancer

Basal-like subtype **+324%**

Triple-negative **+106%**

Luminal B **+74%**

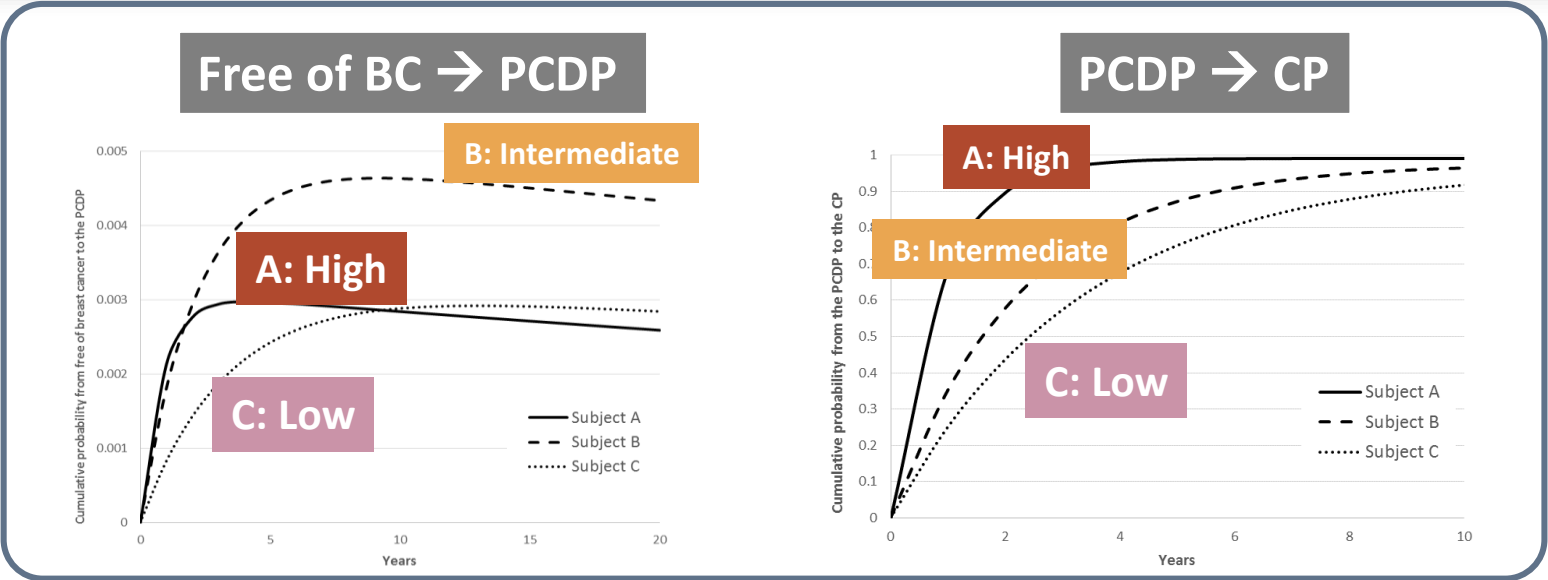
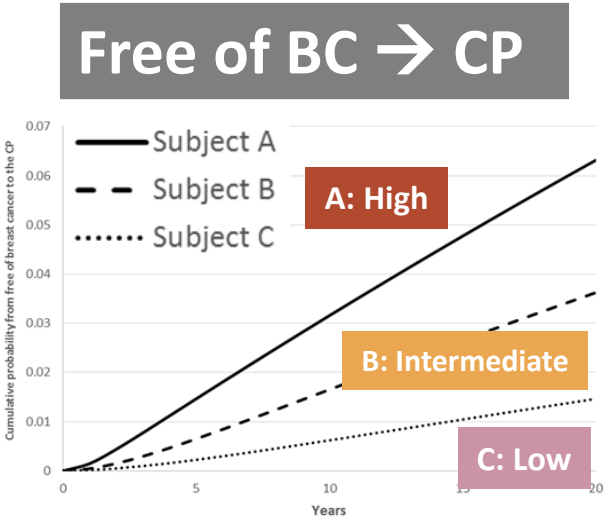
HER-2 positive **+51%**

RISK STRATIFICATION

Empirical data:
Dalarna, Sweden

From Profile to Strategy

Profile	BMI	Late preg.	Dense	Family hx	Subtype	Risk
A	>25	Yes	Yes	Yes	Basal-like	High
B	<25	Yes	Yes	No	HER-2+	Intermediate
C	<25	No	No	No	Luminal A	Low



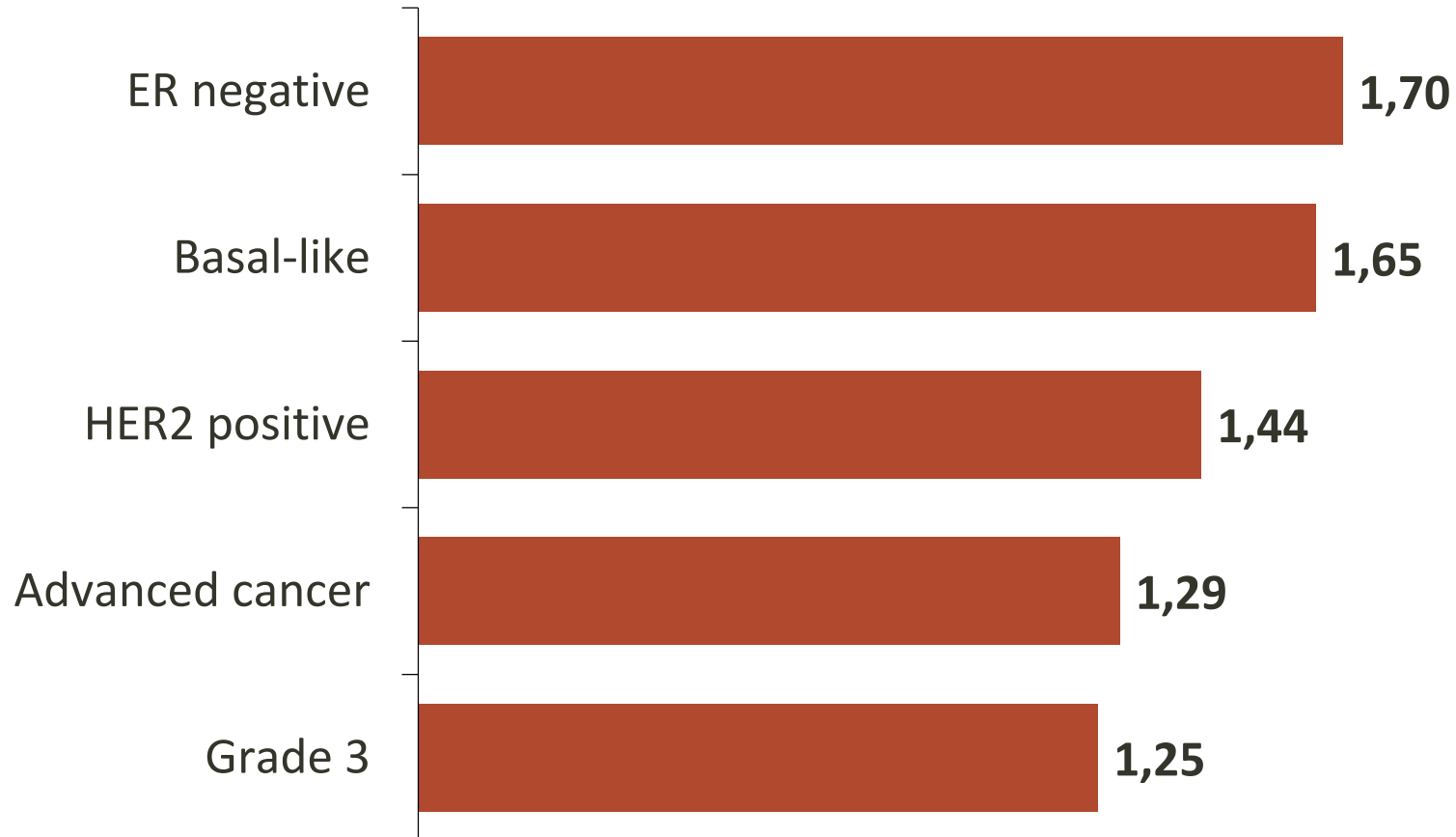
PROGRESSION RISK

Empirical data:
Dalarna, Sweden

Who Leaves the Preclinical Phase Faster?

METAVVERSE

Multivariable hazard ratio of progression PCDP → clinical



HR > 1 = shorter sojourn time

Grade, ER, HER2 and basal-like status are independent drivers ($p < 0.001$)

Mammographic appearance adds little once biology is known ($p = 0.49$)

EXPECTED INTERVAL-CANCER PROPORTION

Empirical data:
Dalarna, Sweden

How Many Cancers Surface Before
the Next Screen?

METAVVERSE

				Expected interval-cancer proportion, % (95% CI)		
		N	MST, years	T = 1 year	T = 2 years	T = 3 years
Grade	1	1,640	4.38	10.6 (10.1–11.1)	19.7 (18.9–20.5)	27.6 (26.5–28.7)
	2	3,175	3.96	11.6 (11.3–12.0)	21.5 (20.9–22.1)	29.9 (29.1–30.7)
	3	1,729	3.21	14.1 (13.5–14.7)	25.6 (24.6–26.6)	35.0 (33.8–36.2)
Mammographic appearance	Stellate/circular	5,296	3.83	12.0 (11.7–12.3)	22.1 (21.6–22.6)	30.7 (30.0–31.3)
	Crushed-stone like	582	4.18	11.1 (10.2–11.9)	20.5 (19.1–22.0)	28.6 (26.8–30.5)
	Powdery	212	4.37	10.6 (9.3–12.0)	19.8 (17.5–22.1)	27.7 (24.7–30.6)
HER2	Negative	3,579	3.86	11.9 (11.5–12.3)	22.0 (21.4–22.6)	30.5 (29.7–31.3)
	Positive	477	2.20	19.6 (18.1–21.2)	34.3 (32.0–36.6)	45.4 (42.8–48.0)
Basal-like	Negative	2,467	4.12	11.2 (10.8–11.6)	20.8 (20.1–21.5)	29.0 (28.1–29.9)
	Positive	239	1.62	25.4 (22.8–28.0)	42.6 (38.9–46.1)	54.5 (50.6–58.1)

PERSONALIZED PROFILES

Empirical data:
Dalarna, Sweden

METAVVERSE

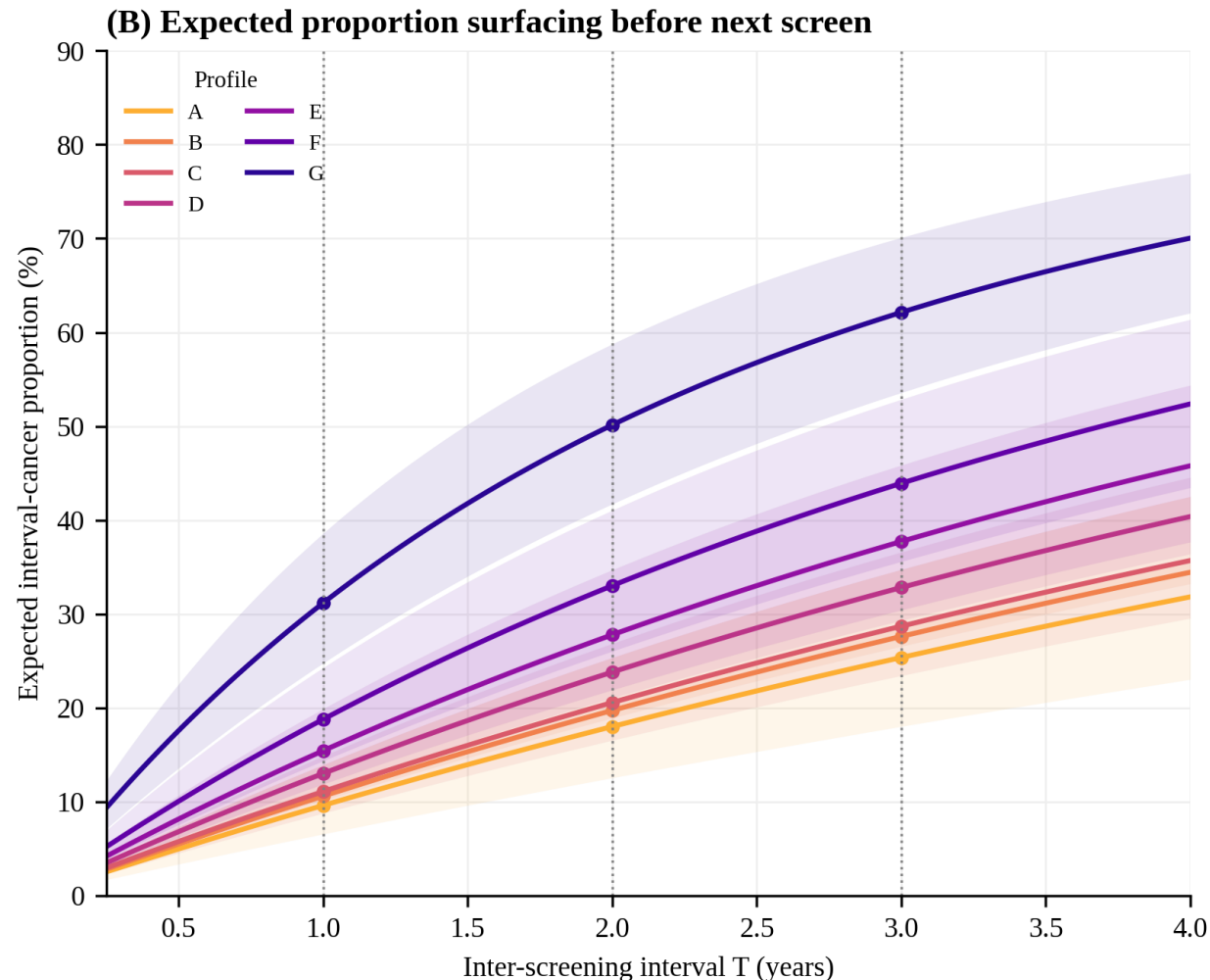
		MST, years	T = 1 y	T = 2 y	T = 3 y
			surfacing before next screen		
A	Grade 1 · HER2– · non-basal · powdery	4.87	9.6%	18.0%	25.4%
B	Grade 1 · HER2– · non-basal · stellate	4.38	10.6%	19.7%	27.6%
C	Grade 2 · HER2– · non-basal · crushed-stone	4.17	11.1%	20.6%	28.7%
D	Grade 3 · HER2– · non-basal · stellate	3.50	13.0%	23.8%	32.8%
E	Grade 2 · HER2+ · non-basal · crushed-stone	2.90	15.4%	27.8%	37.7%
F	Grade 3 · HER2+ · non-basal · crushed-stone	2.32	18.8%	33.0%	43.9%
G	Grade 3 · basal-like (ER–, HER2–) · stellate	1.25	31.2%	50.1%	62.1%

MATCHING THE INTERVAL

The Screening Interval Should Follow the Sojourn Time

Empirical data:
Dalarna, Sweden

METAVVERSE



At a 2-year interval

Profile A: 18% surface
before the next screen

Profile G: 50%

*Fast-progressing biology
needs a shorter interval or
a more sensitive test*

TAILORED PROTOCOL

Screening Matched to Risk

Risk percentile	Start age	Interval	Imaging tool
90–100	30	4.8 months	MRI + Mammography
70–90	35–40	1–1.5 yr	Mammography + Ultrasound
50–70	45–50	2–3 yr	Mammography + Ultrasound
< 50	55–60	4–10 yr	Mammography

Higher risk → *earlier start, tighter interval, more sensitive imaging*

Thank you for your attention

