Scientific cartography

Einar Holsbø and Karina Standahl Olsen

DAGs: mapping causal assumptions

- D: directed

- A: acyclic

- G: graph

DAGs: mapping causal assumptions

- D: directed

- A: acyclic

- G: graph — it's a drawing of points connected by lines.

DAGs: mapping causal assumptions

- D: directed

- A: acyclic

- G: graph — it's a drawing of points connected by lines.

  - Causes/effects are points, a connecting line denotes a causal relationship

DAGs: mapping causal assumptions

- D: directed — a cause points in the direction of its effect

- A: acyclic

- G: graph — it's a drawing of points connected by lines.

  - Causes/effects are points, a connecting line denotes a causal relationship

DAGs: mapping causal assumptions

- D: directed — a cause points in the direction of its effect

- A: acyclic — a phenomenon cannot be its own cause

- G: graph — it's a drawing of points connected by lines.

  - Causes/effects are points, a connecting line denotes a causal relationship

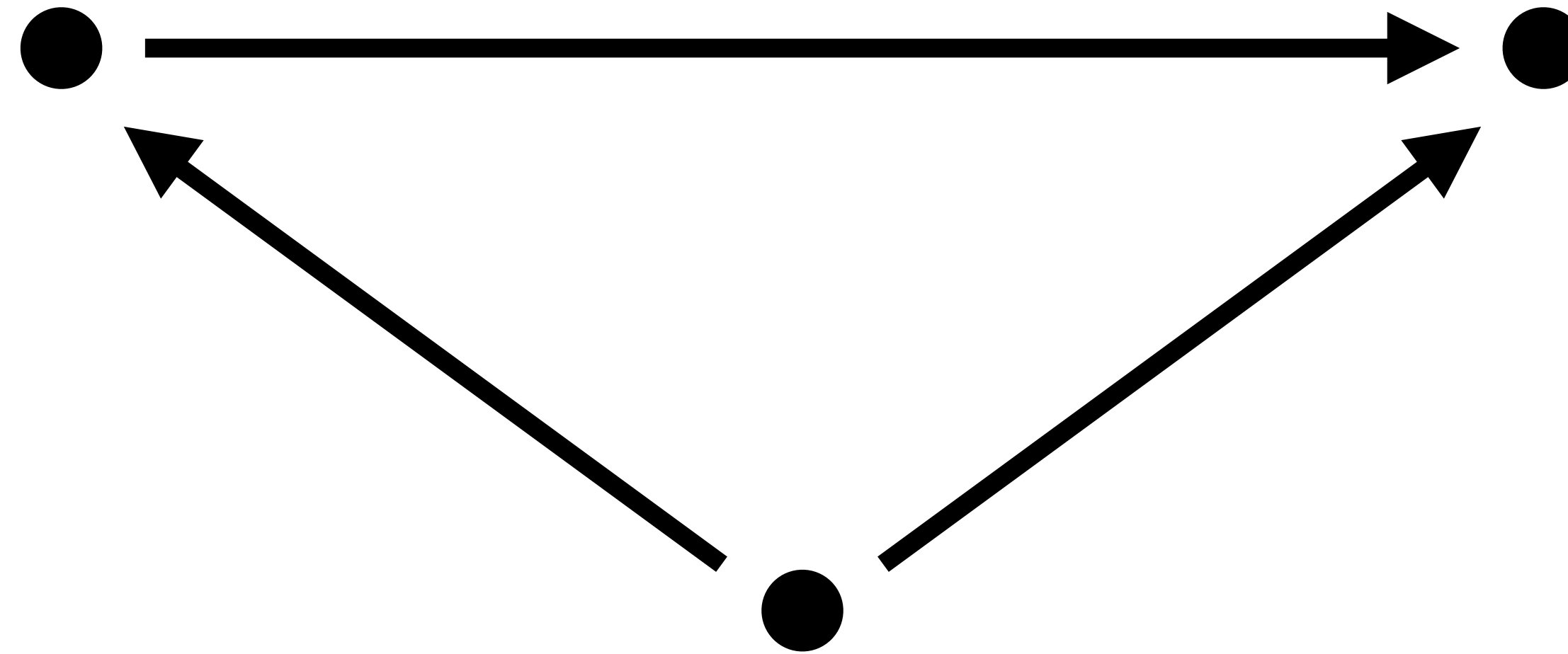# Basic structures

A causal relationship
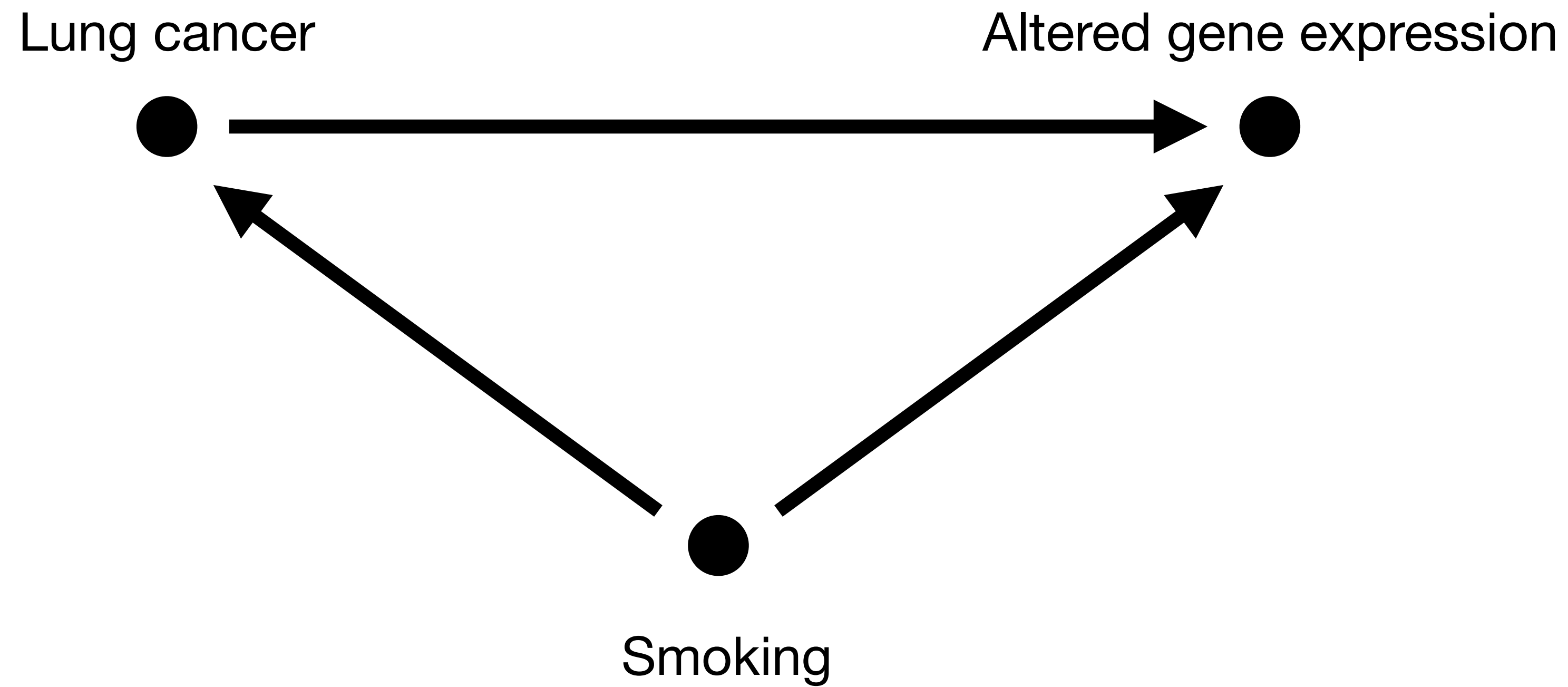
Basic structures

A causal relationship

Smoking → Lung cancer

Basic structures

Confounding: a mixing of effects

Basic structures

Confounding: a mixing of effects

Basic structures

Confounding: a mixing of effects

Lung cancer

Altered gene expression
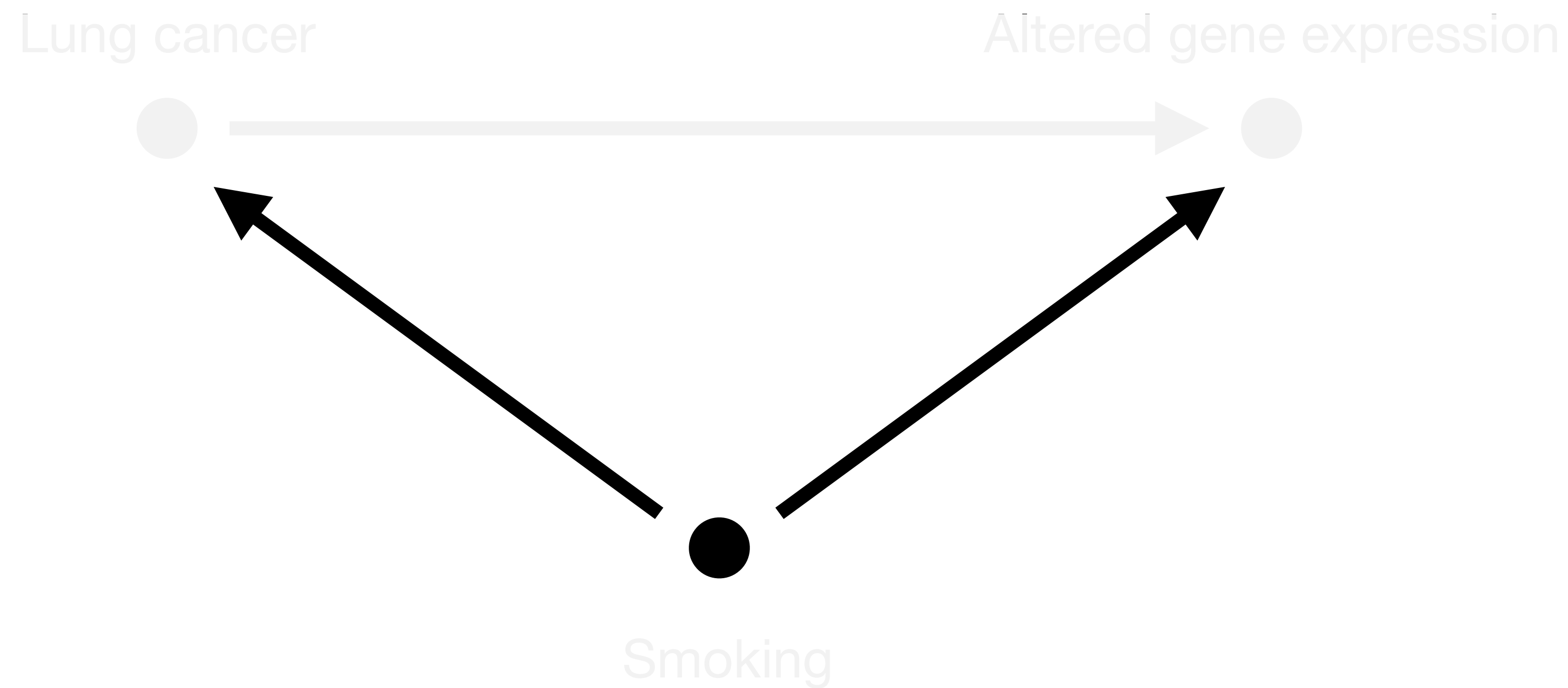
Smoking

"Back-door path:" closed by controlling for smoking

Basic structures
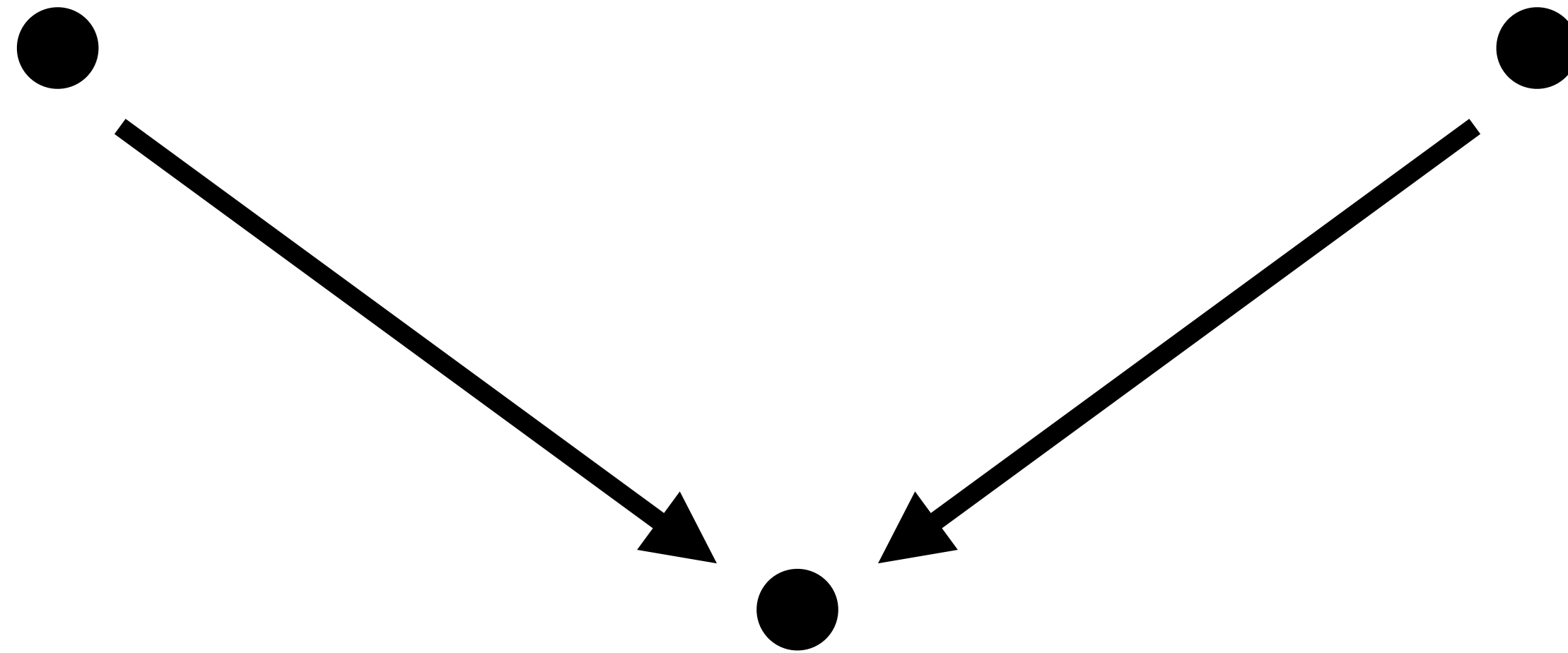
Collider: a closed path

Basic structures

Collider: a closed path

Smoking

Asbestos work

Lung cancer

Basic structures

Collider: a closed path

Smoking   Asbestos work

**Conditioning on lung cancer introduces spurious association:
the one "explains away" the other**

Lung cancer

Basic structures

Collider: a closed path

Respiratory disease

Bone disease

Hospitalization
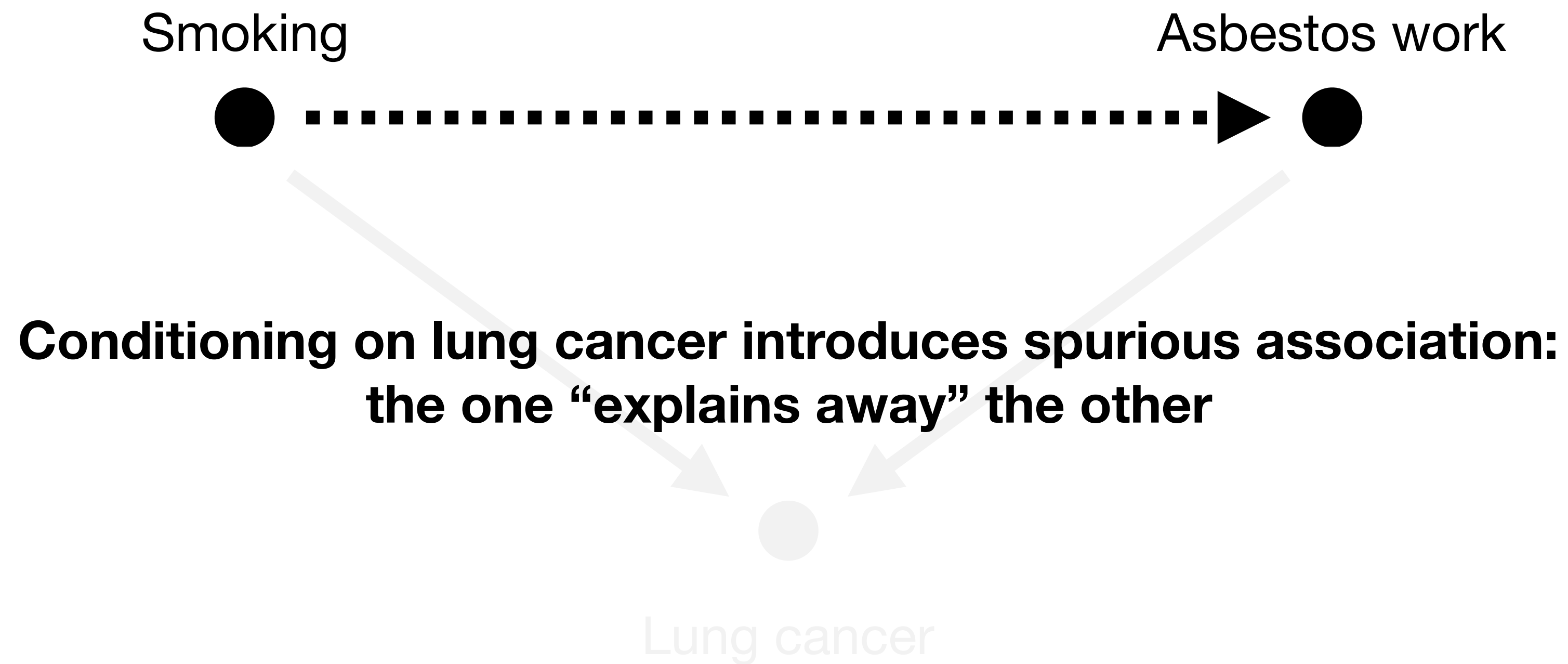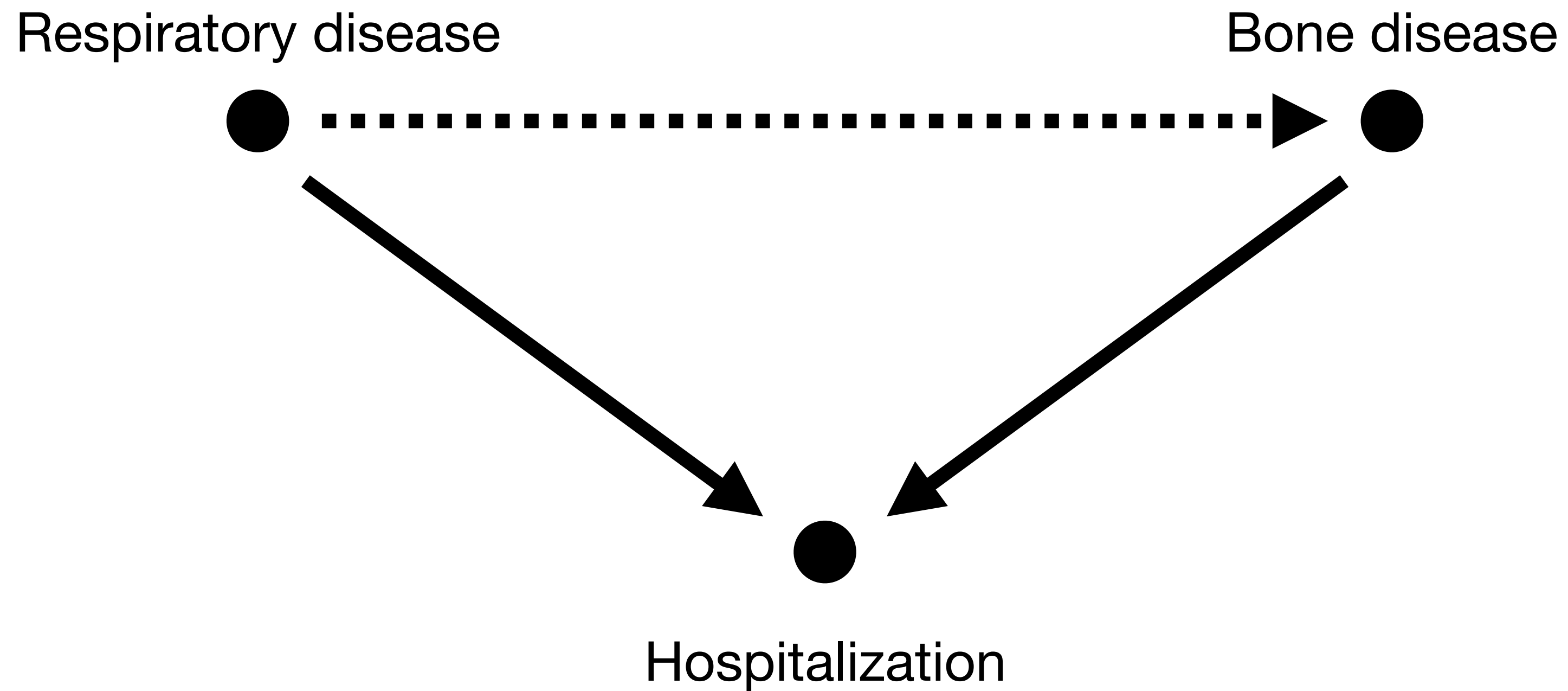
Richard McElreath
@rlmcelreath

The Collider: A Weekly Illustrated Journal of Science

3. *The Collider*

PUBLISHED IN NATURE

NEWSWORTHY   RIGOROUS

9:45 AM · Oct 29, 2019 · Twitter Web App

Mapmaking

- Draw causal graph using plausible/suspected causal relationships

- Can get complicated

- Use rules of "d-separation" to ascertain what to adjust for (ie. put the graph into a computer).

UNDERGROUND

& Latimer
Rickmansworth
Moor Park
Northwood
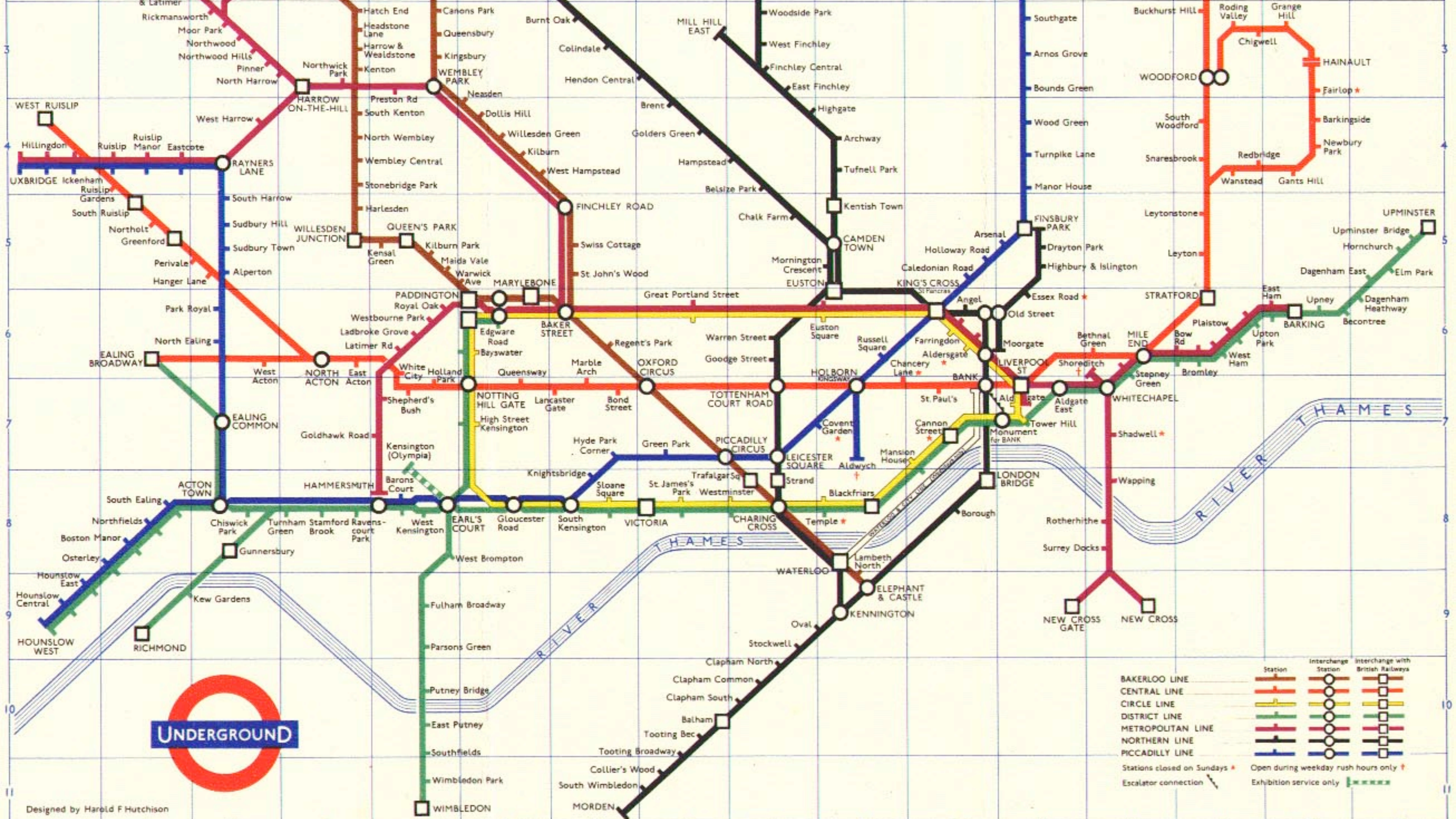Northwood Hills
Pinner
North Harrow
Hatch End
Headstone Lane
Harrow & Wealdstone
Kenton
Canons Park
Queensbury
Kingsbury
Burnt Oak
Woodside Park
West Finchley
MILL HILL EAST
Colindale
Finchley Central
East Finchley
Southgate
Arnos Grove
Bounds Green
Wood Green
Buckhurst Hill
Roding Valley
Grange Hill
Chigwell
HAINAULT
Fairlop *
Barkingside
Newbury Park
WOODFORD
Hendon Central
Brent
Highgate
WEMBLEY PARK
Neasden
Preston Rd
South Kenton
North Wembley
Wembley Central
Stonebridge Park
Harlesden
Northwick Park
HARROW ON-THE-HILL
West Harrow
Dollis Hill
Willesden Green
Kilburn
West Hampstead
Golders Green
Hampstead
Archway
Tufnell Park
Turnpike Lane
Manor House
South Woodford
Snaresbrook
Redbridge
Wanstead
Gants Hill
WEST RUISLIP
Hillingdon Ickenham
UXBRIDGE
Ruislip Gardens
South Ruislip
Northolt
Greenford
Perivale
Hanger Lane
Ruislip Manor
Ruislip
Eastcote
RAYNERS LANE
South Harrow
Sudbury Hill
Sudbury Town
Alperton
WILLESDEN JUNCTION
QUEEN'S PARK
Kensal Green
Kilburn Park
Maida Vale
Warwick Ave
FINCHLEY ROAD
Swiss Cottage
St John's Wood
Chalk Farm
Belsize Park
Kentish Town
CAMDEN TOWN
Mornington Crescent
EUSTON
Arsenal
Holloway Road
Caledonian Road
FINSBURY PARK
Drayton Park
Highbury & Islington
Leytonstone
Leyton
UPMINSTER
Upminster Bridge
Hornchurch
Dagenham East
Elm Park
Park Royal
North Ealing
West Acton
EALING BROADWAY
NORTH ACTON
East Acton
PADDINGTON
Royal Oak
Westbourne Park
Ladbroke Grove
Latimer Rd
White City
Shepherd's Bush
Goldhawk Road
Edgware Road
Bayswater
MARYLEBONE
BAKER STREET
Great Portland Street
Regent's Park
Warren Street
Goodge Street
Euston Square
Russell Square
KING'S CROSS
St Pancras
Angel
Essex Road *
Old Street
Farringdon
Aldersgate
Moorgate
STRATFORD
East Ham
Upton Park
Plaistow
BARKING
Bethnal Green
MILE END
Bow Rd
West Ham
Becontree
Dagenham Heathway
Upney
Holland Park
Notting Hill Gate
Queensway
Lancaster Gate
Marble Arch
Bond Street
OXFORD CIRCUS
TOTTENHAM COURT ROAD
HOLBORN
Chancery Lane *
St. Paul's
BANK
LIVERPOOL ST
Shoreditch †
Stepney Green
Bromley
High Street Kensington
Hyde Park Corner
Green Park
PICCADILLY CIRCUS
LEICESTER SQUARE
Covent Garden
Cannon Street
Mansion House
Monument for BANK
Aldgate
Aldgate East
Tower Hill
WHITECHAPEL
Shadwell *
EALING COMMON
Knightsbridge
Sloane Square
St. James's Park
Trafalgar Sq
Strand
Aldwych
ACTON TOWN
HAMMERSMITH
Barons Court
West Kensington
EARL'S COURT
Gloucester Road
South Kensington
VICTORIA
Westminster
CHARING CROSS
Blackfriars
Temple *
LONDON BRIDGE
Borough
Wapping
Rotherhithe
Surrey Docks
South Ealing
Northfields
Boston Manor
Osterley
Hounslow East
Hounslow Central
Chiswick Park
Turnham Green
Stamford Brook
Ravenscourt Park
West Brompton
Kensington (Olympia)
Gunnersbury
Kew Gardens
Fulham Broadway
WATERLOO
Lambeth North
ELEPHANT & CASTLE
KENNINGTON
NEW CROSS GATE
NEW CROSS
HOUNSLOW WEST
RICHMOND
Parsons Green
Putney Bridge
East Putney
Southfields
WIMBLEDON
Wimbledon Park
Oval
Stockwell
Clapham North
Clapham Common
Clapham South
Balham
Tooting Bec
Tooting Broadway
Colliers Wood
South Wimbledon
MORDEN

RIVER THAMES

Legend

BAKERLOO LINE
CENTRAL LINE
CIRCLE LINE
DISTRICT LINE
METROPOLITAN LINE
NORTHERN LINE
PICCADILLY LINE

Station   Interchange Station   Interchange with British Railways

Stations closed on Sundays *
Escalator connection
Open during weekday rush hours only †
Exhibition service only

Designed by Harold F Hutchison

On Exactitude in Science

Jorge Luis Borges, *Collected Fictions*, translated by Andrew Hurley.

…In that Empire, the Art of Cartography attained such Perfection that the map of a single Province occupied the entirety of a City, and the map of the Empire, the entirety of a Province. In time, those Unconscionable Maps no longer satisfied, and the Cartographers Guilds struck a Map of the Empire whose size was that of the Empire, and which coincided point for point with it. The following Generations, who were not so fond of the Study of Cartography as their Forebears had been, saw that that vast Map was Useless, and not without some Pitilessness was it, that they delivered it up to the Inclemencies of Sun and Winters. In the Deserts of the West, still today, there are Tattered Ruins of that Map, inhabited by Animals and Beggars; in all the Land there is no other Relic of the Disciplines of Geography.

—Suarez Miranda,*Viajes de varones prudentes*, Libro IV,Cap. XLV, Lerida, 1658

Nøkkeltall

Nøkkeltall	
Nye tilfeller i 2015	3 415
Antall nye tilfeller per 100 000 i 2015	128.0
Høyest insidens i aldersgruppen	65–69 år
Akkumulert risiko fram til 75 års alder	8.5 %
Fem-års relativ overlevelse 2011–15	89.0 %

Brystkreft
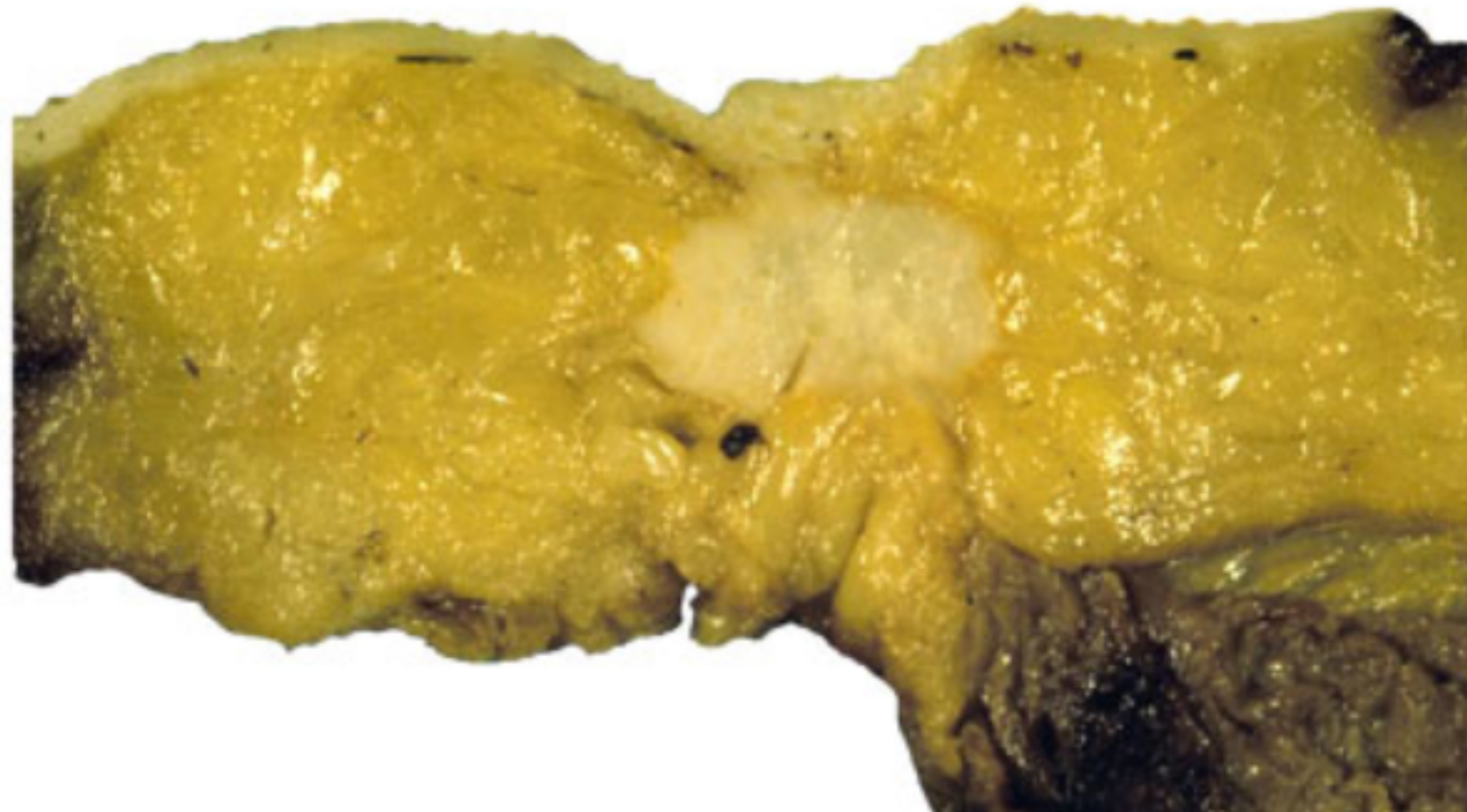### Insidensrater i ulike aldersgrupper
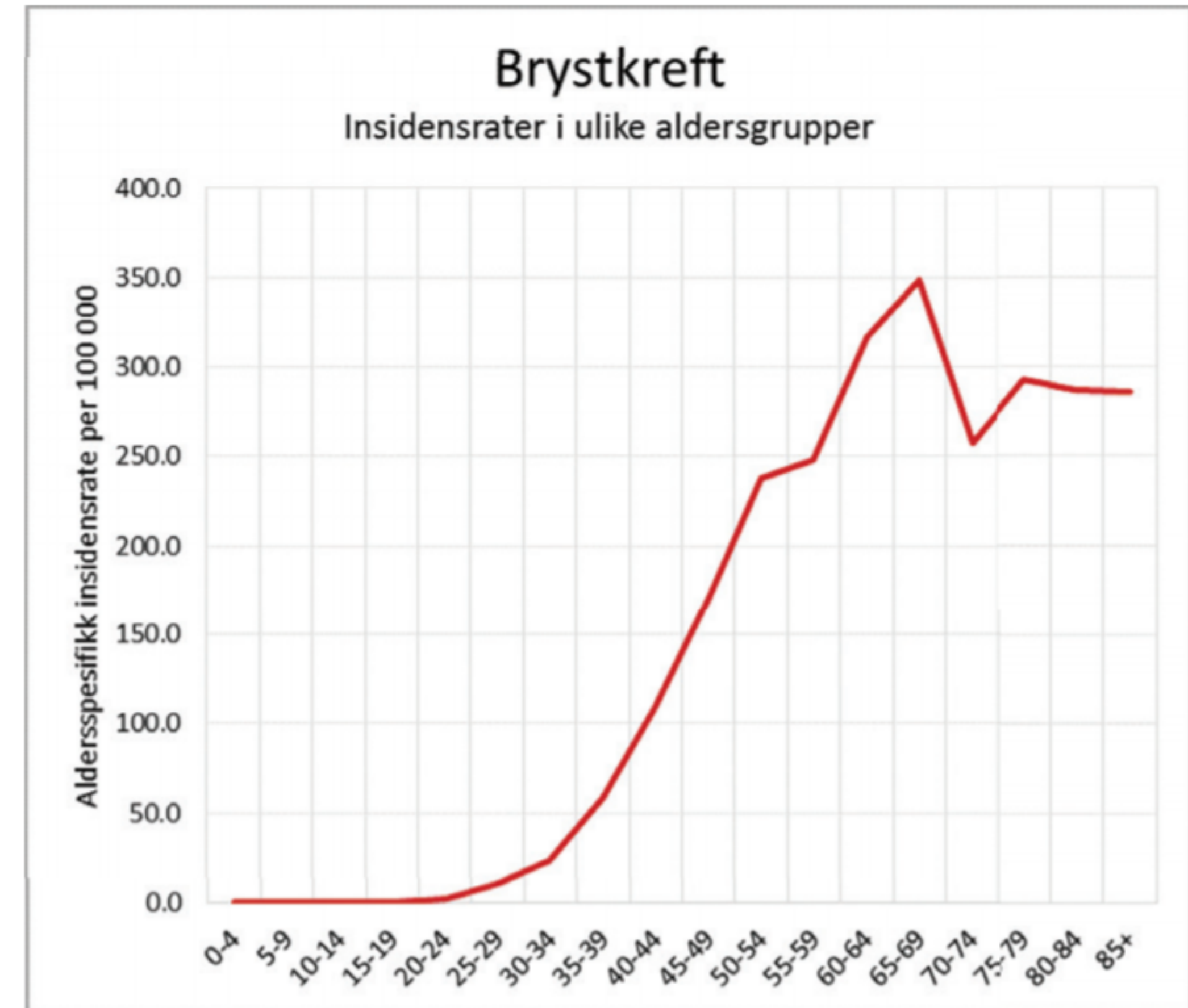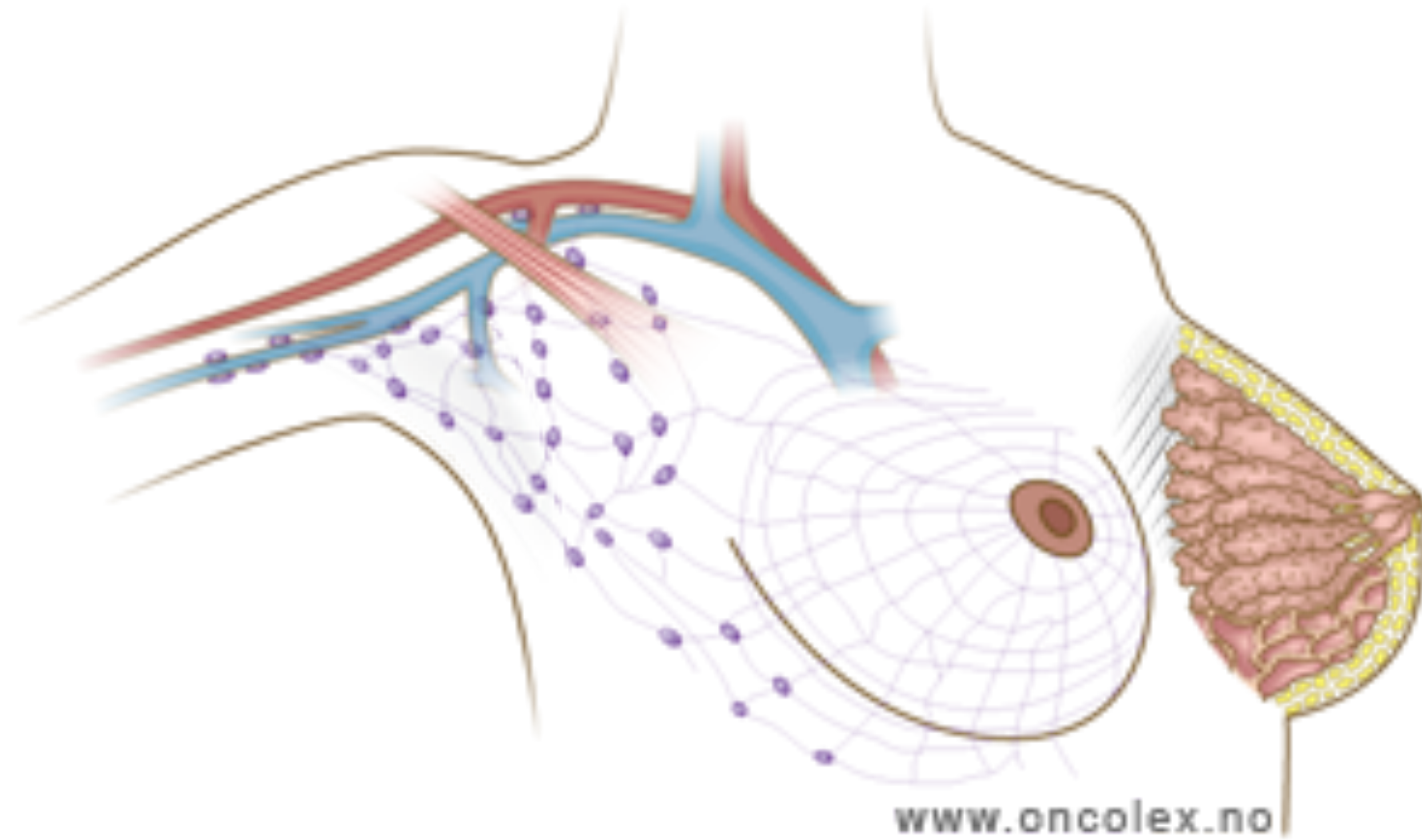
Classification and staging

T - Tumor size and placement
N - Cancer cells in adjacent lymph node
M - Distant metastases

TNM dictates stage, I-IV
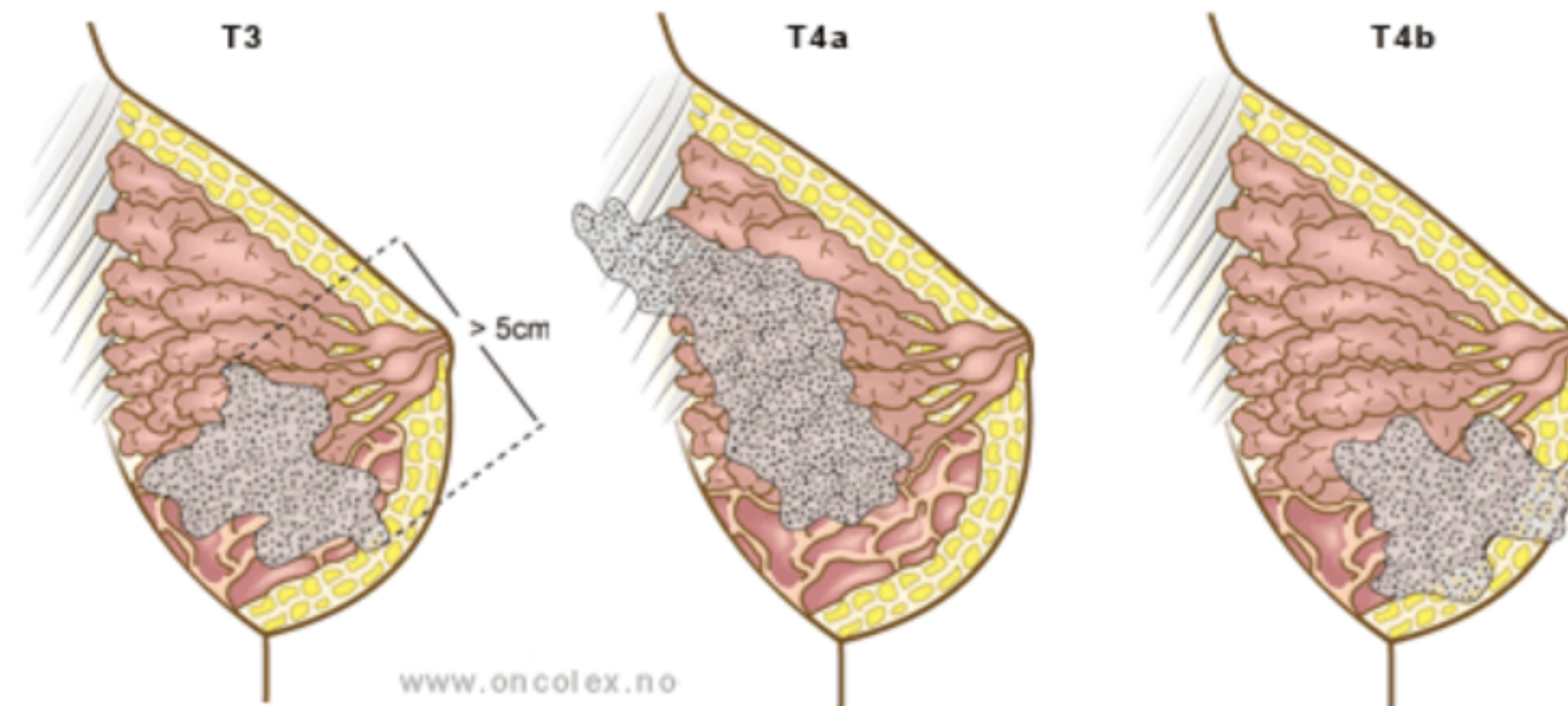
Tumor (3-4 cm) and fatty tissue

FISH analysis to determine Her2

Presence of certain proteins...

ER – estrogen receptor
PR – progesterone receptor
Her2 – Human epidermal growht factor receptor 2

... dictate subroups:
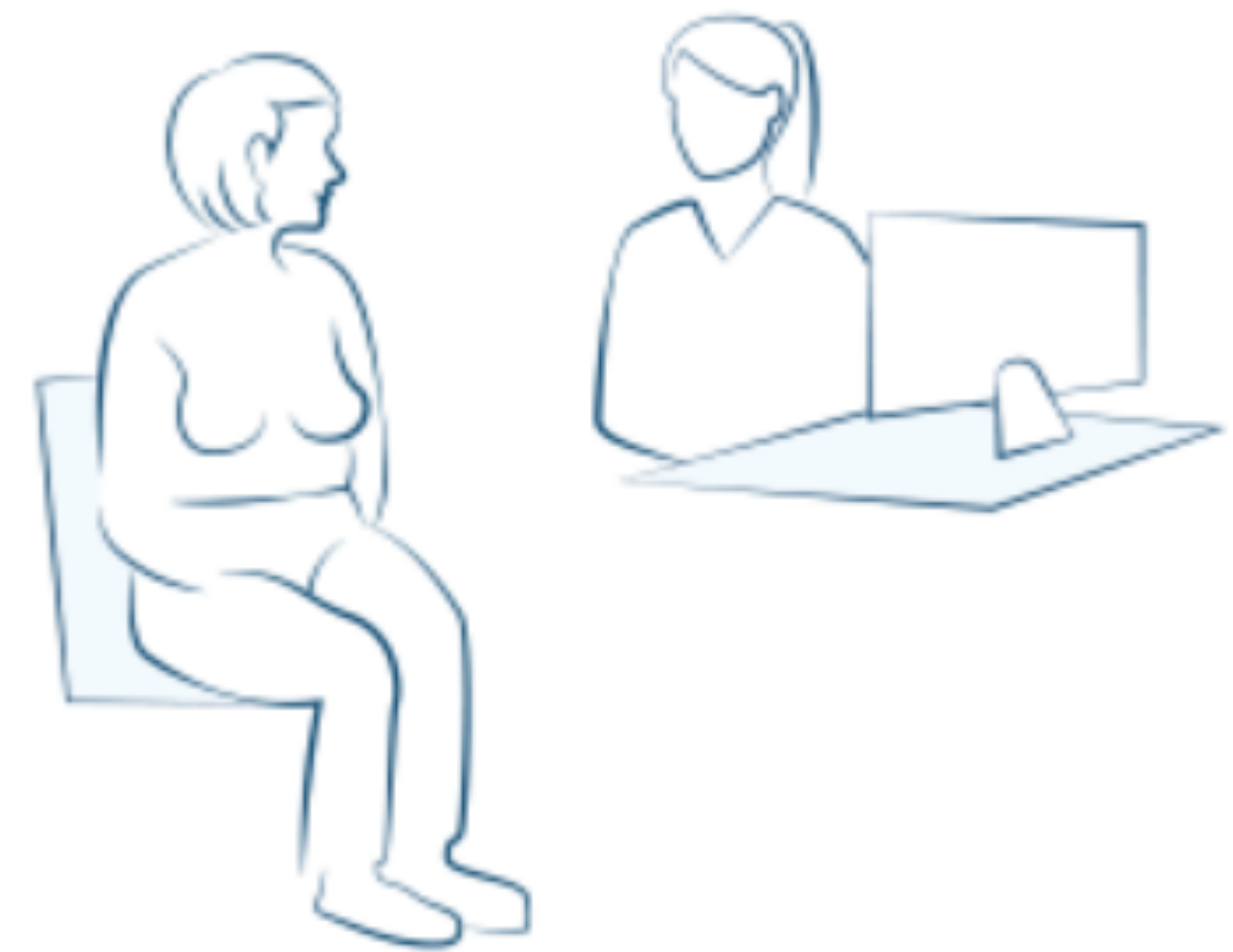
Luminal A
Luminal B
Her2 positive
Tripple negative

Mammography screening
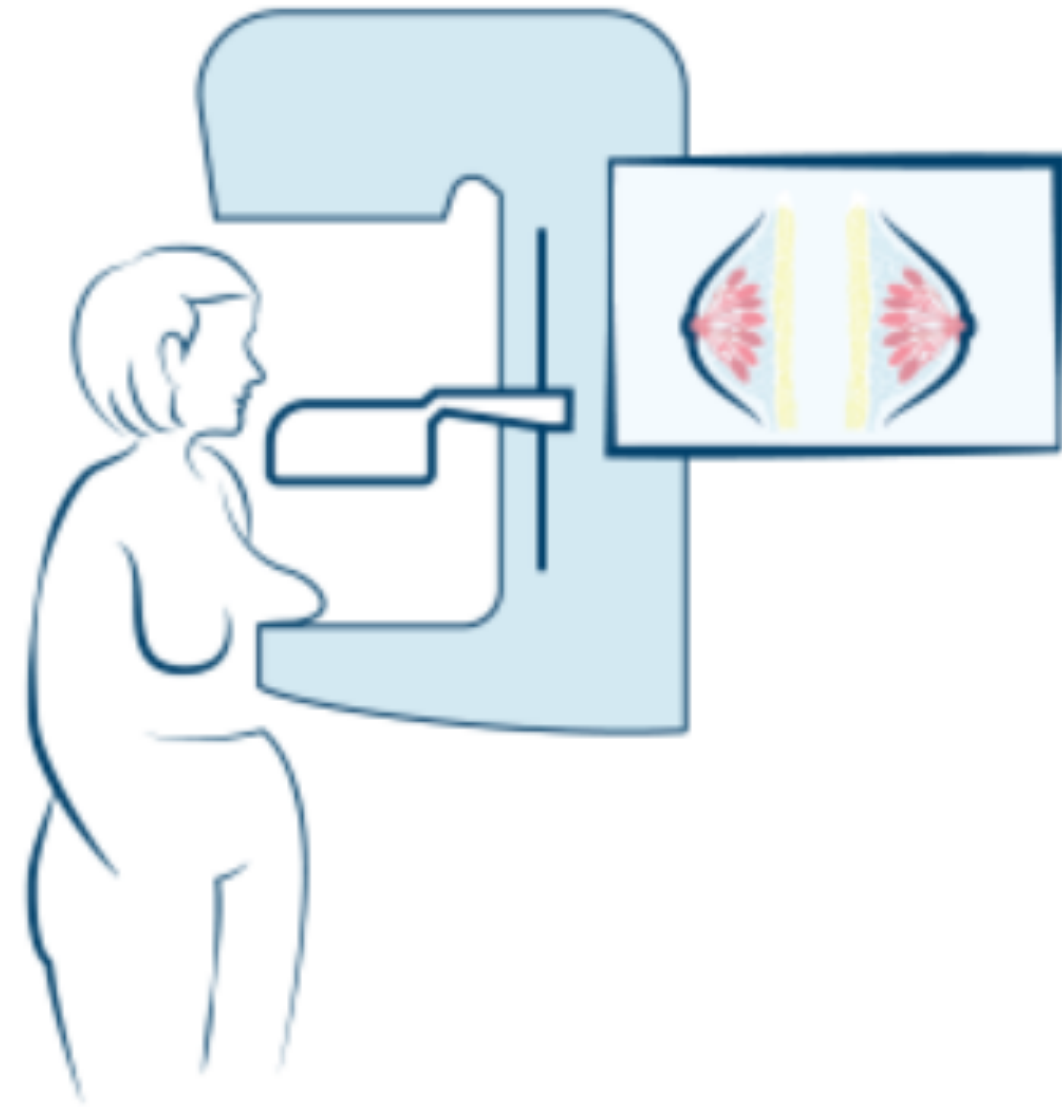
Clinically detected

Search Patterns
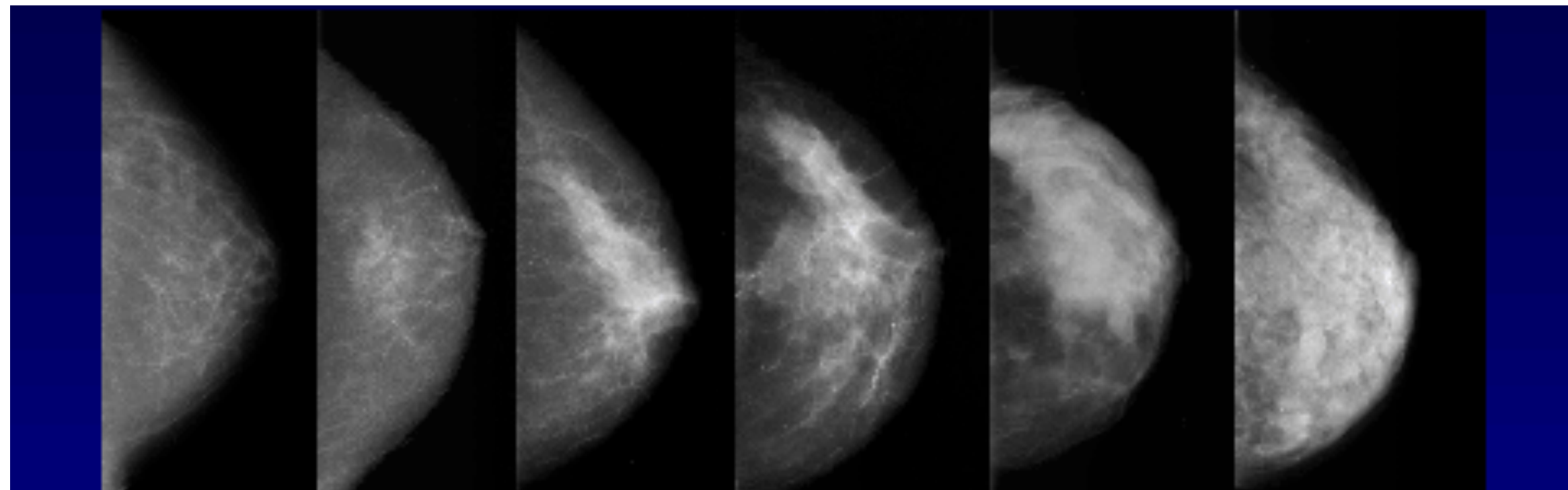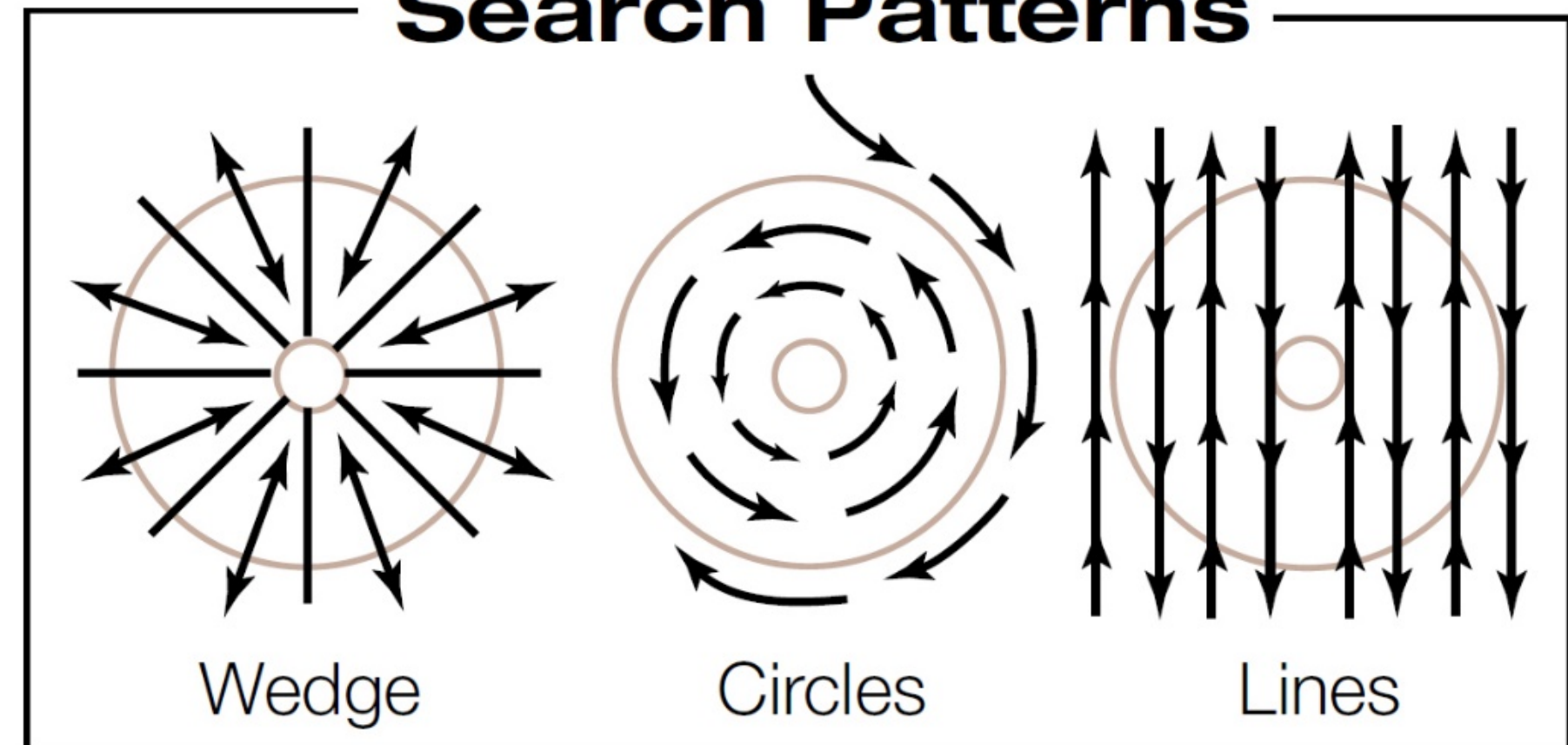
Wedge   Circles   Lines

Risk factors:

Sex, age

Hereditability: BRCA genes, others

Hormone load, endo- and exogenous
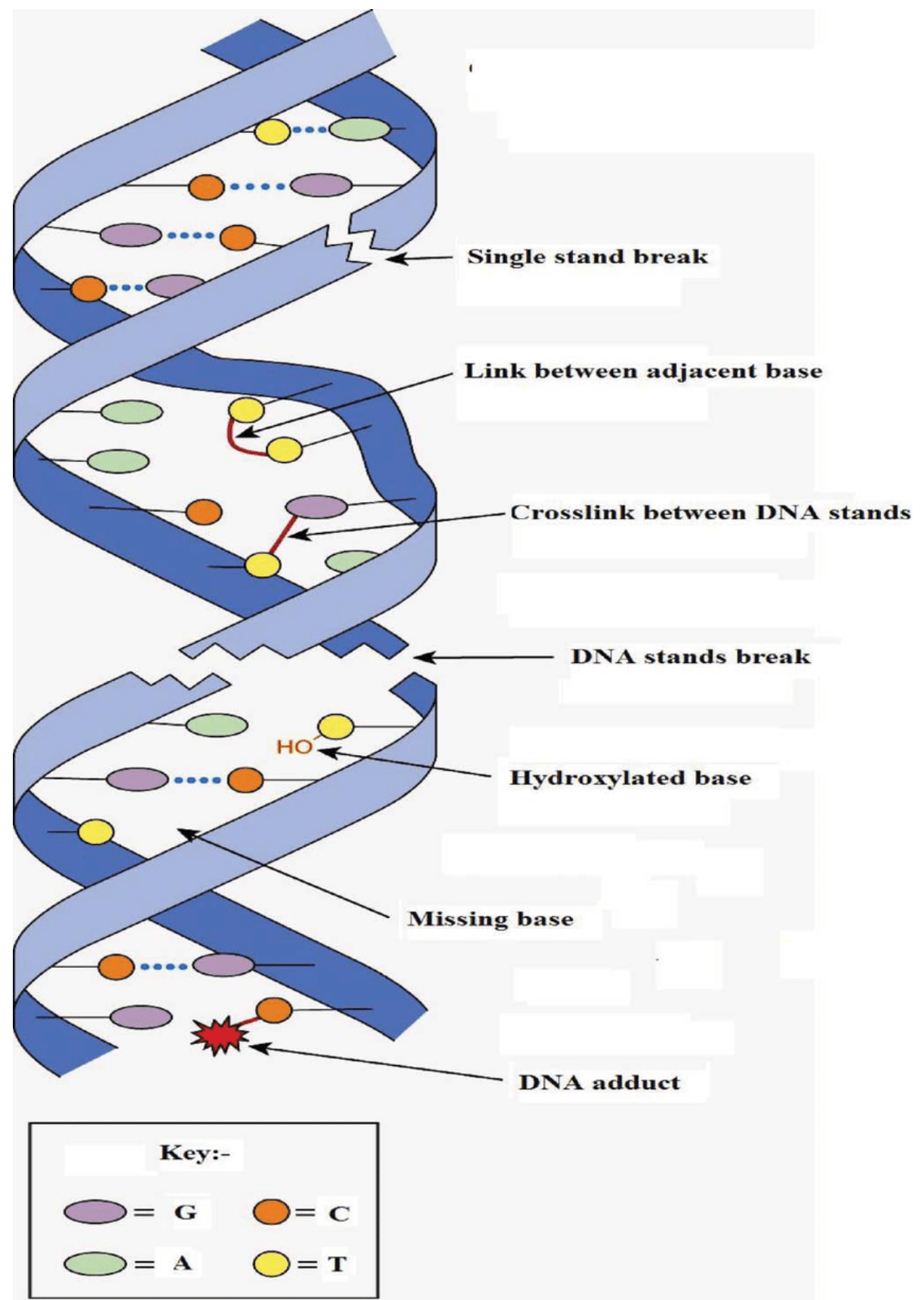
Overweight/obesity

Height

Alcohol use

Smoking

Protective factors:

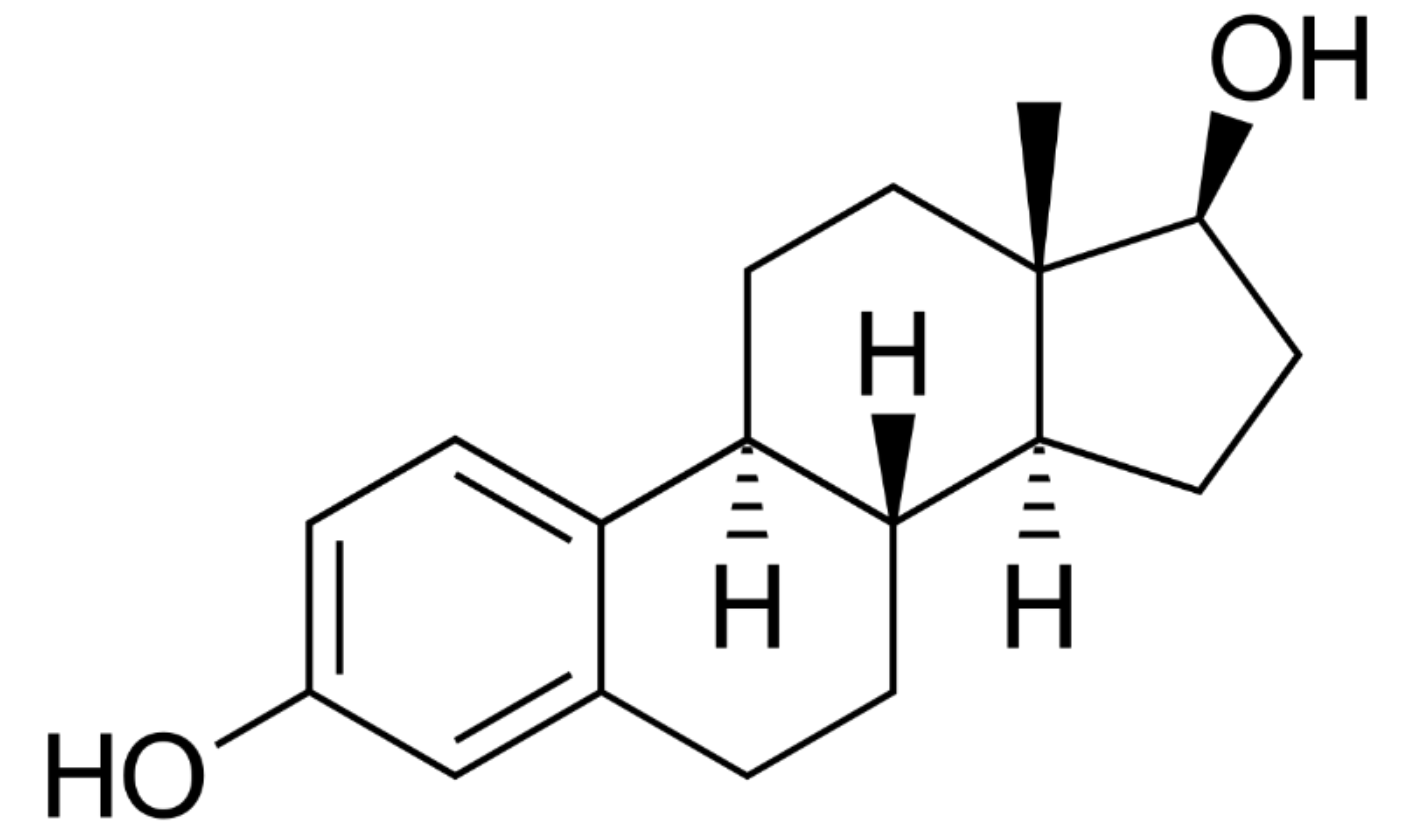Pregnancies: early and many

Breast feeding

Asian origin

Physical activity

Key:-
= G   = C
= A   = T

DNA damage ->

Uncontrolled cell proliferation ->

Cancer

Estradiol

Hormones act as mitogens ->

Increased cell proliferation ->

Cancer

Risk factors:

Sex, **age**

Hereditability: BRCA genes, others

Hormone load, endo- and exogenous

Overweight/obesity

Height

Alcohol use

Smoking

Protective factors:

Pregnancies: early and many

Breast feeding

Asian origin

Physical activity
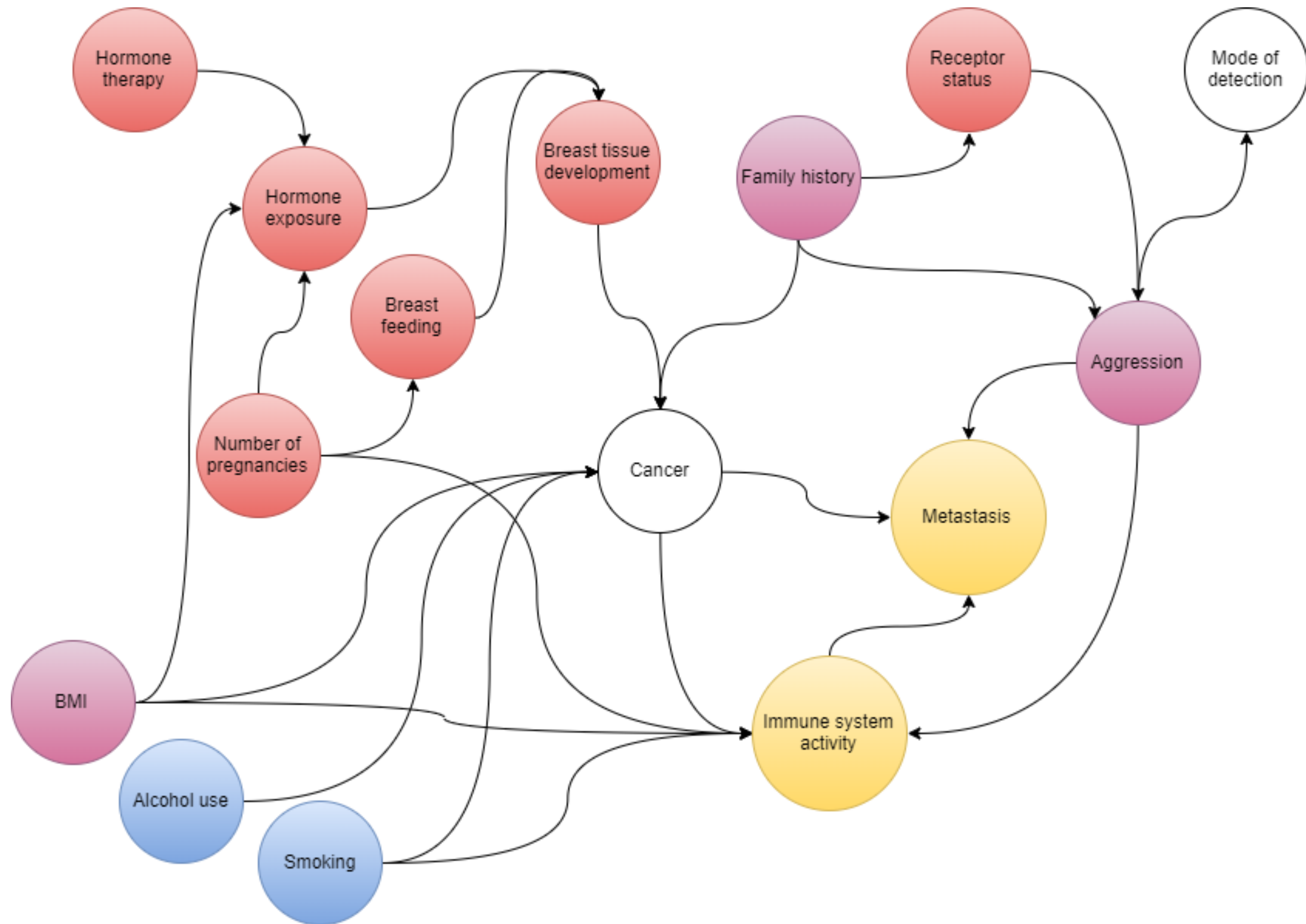
Table 1 Descriptive characteristics of breast cancer cases and healthy controls

	Controls	Breast cancer, non-metastasized	Breast cancer, metastasized
n	156	115	41
Age	56.1	56.0	56.2
BMI	25.5	25.5	26.1
Smoking	37 (24%)	26 (23%)	10 (24%)
HT use	29 (19%)	41 (36%)	12 (29%)
Parity	1.9	1.8	1.8

Table 2 Characteristics of the breast cancer cases

	Non-metastasized	Metastasized
Follow-up time	318	375
Detection mode		
Screening	91 (79%)	20 (49%)
Interval	24 (21%)	21 (51%)
Subtypes		
Luminal A	59 (51%)	26 (63%)
Luminal B	9 (8%)	4 (10%)
Triple negative	2 (2%)	3 (7%)
HER2 positive	0	3 (7%)
Unknown	45 (39%)	5 (13%)

```r
library(ggdag)
theme_set(theme_dag())
metastasis_dag <- dagify(
  cancer ~ smoking + alcohol + bmi + tissue_development + family_history,
  metastasis ~ cancer + aggression + immune_activity,
  detection ~ aggression,
  tissue_development ~ breast_feeding + hormone_exposure,
  hormone_exposure ~ pregnancies + bmi + hormone_therapy,
  breast_feeding ~ pregnancies,
  immune_activity ~ pregnancies + cancer + bmi + smoking + aggression,
  aggression ~ family_history + receptor_status,
  receptor_status ~ family_history,

  exposure = "immune_activity",
  outcome ="metastasis"
)

ggdag_adjustment_set(metastasis_dag, text_col = "black")
```

{aggression, cancer}

Thanks!