Karolinska Institutet

Natural History Models for Breast Cancer Growth and Spread

Alessandro Gasparini *and* Keith Humphreys

Department of Medical Epidemiology and Biostatistics

Karolinska Institutet, Stockholm, Sweden

YSM 2020 · @ellessenne · alessandro.gasparini@ki.se

The Take-Home Messages

1. Continuous growth model can be used to model the unobserved natural history of a tumour. They are more efficient than Markov models when analysing several factors at once;

The Take-Home Messages

1. Continuous growth model can be used to model the unobserved natural history of a tumour. They are more efficient than Markov models when analysing several factors at once;
2. We introduce a model that relates the unobserved past natural history of tumours to the risk of future outcomes, e.g. distant metastatic spread. To the best of our knowledge, this is the first of its kind for breast cancer;

The Take-Home Messages

1. Continuous growth model can be used to model the unobserved natural history of a tumour. They are more efficient than Markov models when analysing several factors at once;
2. We introduce a model that relates the unobserved past natural history of tumours to the risk of future outcomes, e.g. distant metastatic spread. To the best of our knowledge, this is the first of its kind for breast cancer;
3. This is fundamentally important to assess new interventions and treatments e.g. in a microsimulation framework.

Breast Cancer

> *Breast cancer is when abnormal cells in the breast begin to grow and divide in an uncontrolled way and eventually form a growth (tumour).*

It is the most common cancer in the UK, and the most common cancer among women worldwide.

Breast Cancer in Numbers

- 2.09 million cases worldwide (2018)
- ~55K cases in the UK (2016) and ~8K in Sweden (2017)
- 627K deaths from breast cancer worldwide, 5th most common cancer death (2018)
- ~11K deaths from breast cancer in the UK (2015–2017) and ~1.4K in Sweden (2017)

In most cases, breast cancer death is caused by metastases that have spread throughout the body.

Data source: WHO, CRUK, Socialstyrelsen.

Traditional Statistical Approaches

Continuous tumour progression models:

- In the presence of a screening programme, e.g. Weedon-Fekjær et al. (2008, 2010);
- In the absence of a screening programme, e.g. Plevritis et al. (2007);

Multi-state Markov models with three states, e.g. in Duffy et al. (2001):

1. No detectable cancer;
2. Preclinical cancer (only detectable by screening);
3. Clinical symptomatic breast cancer.

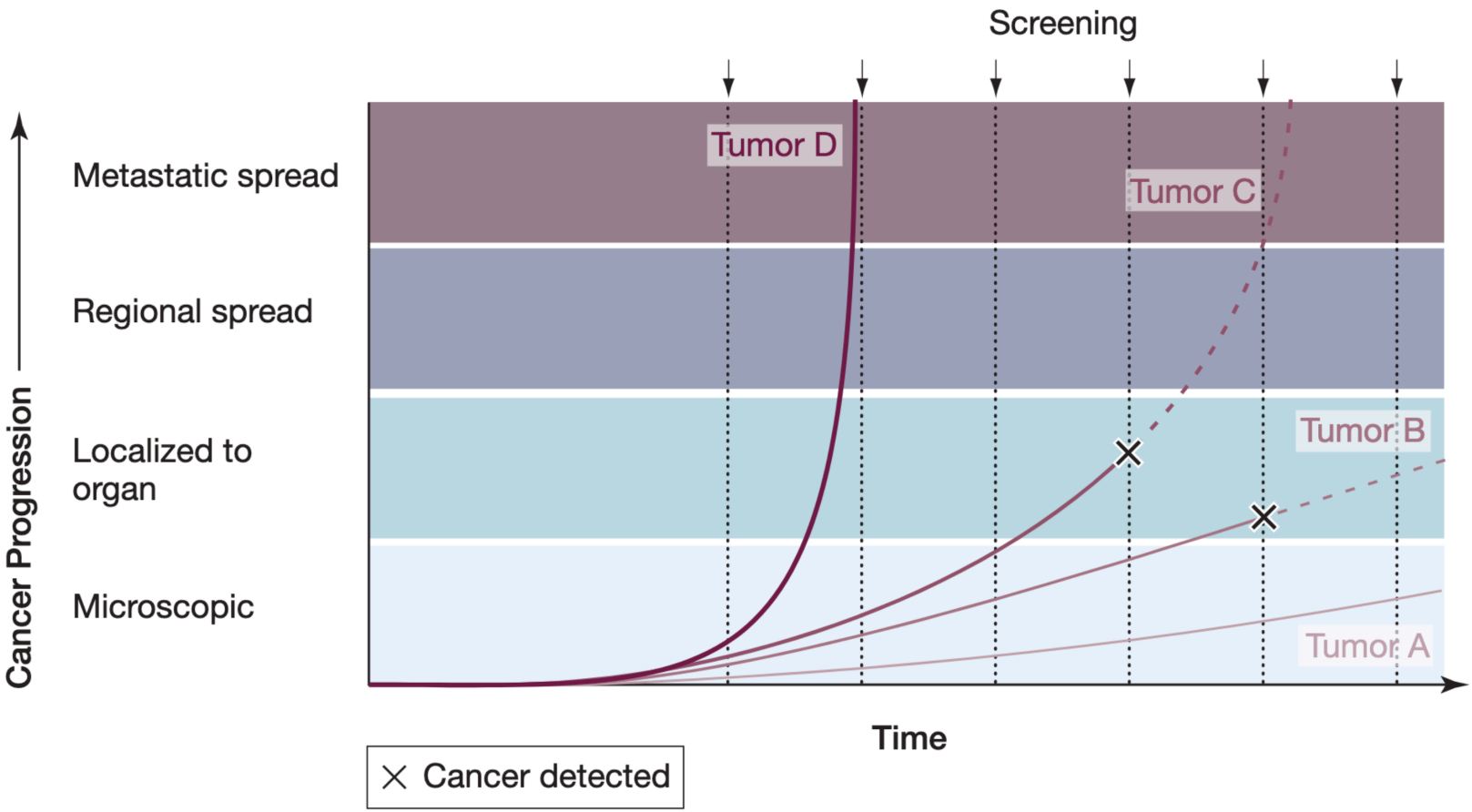
Aims of Our Work

Lots of researchers are looking into individualised screening and understanding how interventions can be tailored to patients' characteristics. The goal is to improve future outcomes.

However, traditional prognostic models only use information that is observed at diagnosis such as tumour size and detection mode.

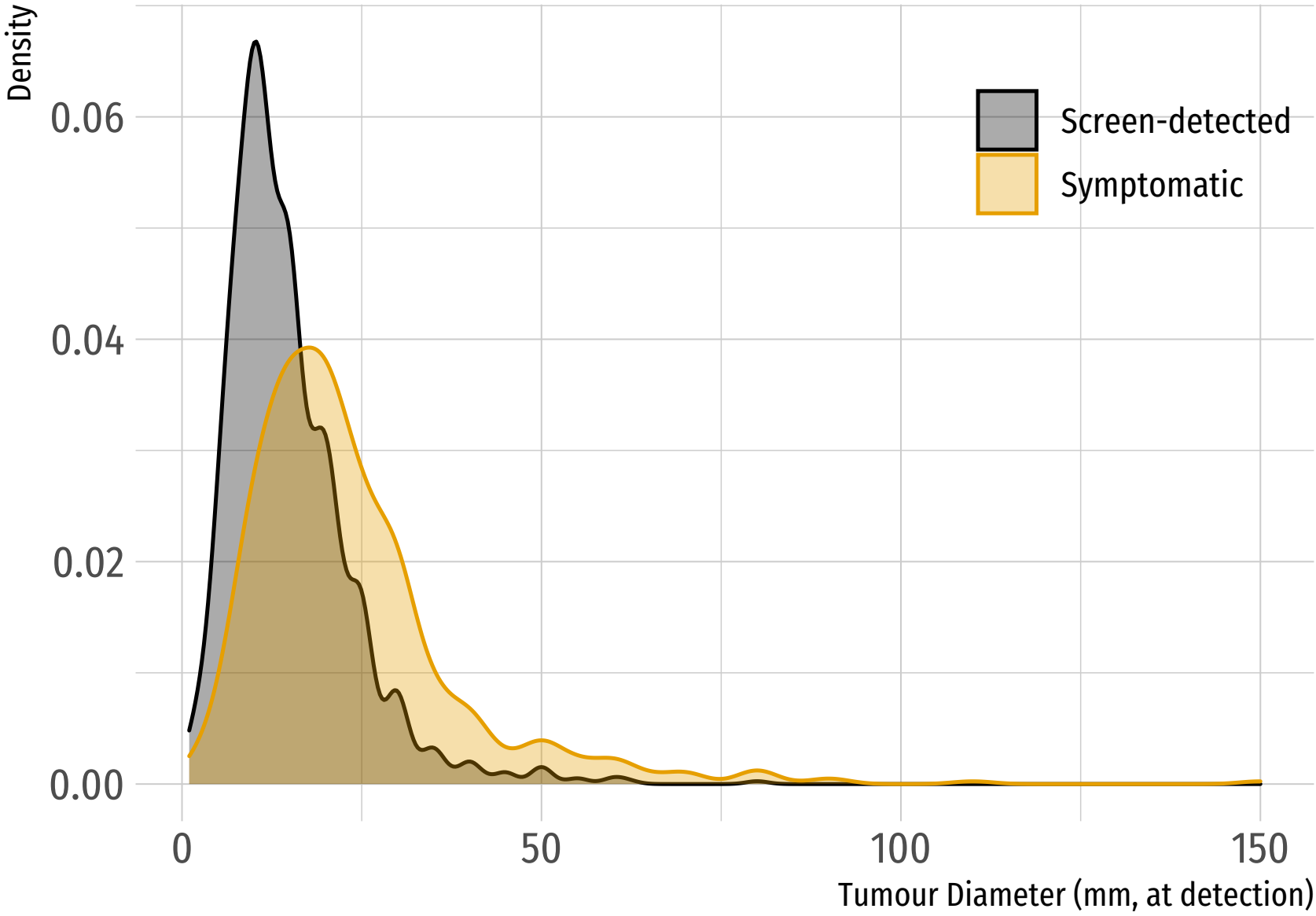
Our aim is to **develop a model that relates the unobserved natural history of the tumour to future patients' outcomes**. For instance, here we study time to distant metastasis.

The Natural History of Tumours is Unobserved



Source: https://doi.org/10.1001/jama.2009.1498

Screen-Detected and Symptomatic Tumours Are Different



Continuous Tumour Growth Models

The tumour volume at time t is assumed to grow exponentially:

$$V(t, r) = V_{Cell} \exp(t/r), \ \forall \ t \geq 0$$

Continuous Tumour Growth Models

The tumour volume at time t is assumed to grow exponentially:

$$V(t, r) = V_{Cell} \exp(t/r), \ \forall \ t \geq 0$$

The inverse growth rates r are assumed to follow a Gamma distribution with shape parameter τ_1 and inverse scale (e.g. rate) parameter τ_2:

$$f_R(r) = \frac{\tau_2^{\tau_1}}{\Gamma(\tau_1)} r^{\tau_1 - 1} \exp(-\tau_2 r)$$

Continuous Tumour Growth Models

On top of the growth functions, we assume that the probability of symptomatic detection at time $t_{det} \geq t$ depends linearly on the size of the tumour (volume):

$$P(t_{det} \in [t, t + dt) | t_{det} \geq t, R = r) = \eta V(t, r)dt + O(dt)$$

Continuous Tumour Growth Models

On top of the growth functions, we assume that the probability of symptomatic detection at time $t_{det} \geq t$ depends linearly on the size of the tumour (volume):

$$P(t_{det} \in [t, t + dt) | t_{det} \geq t, R = r) = \eta V(t, r) dt + O(dt)$$

Finally, screening sensitivity is assumed to follow a *logistic* function:

$$S(X|\beta) = \frac{\exp(X\beta)}{1 + \exp(X\beta)}$$

X can include any covariates that we observe.

Likelihood Function

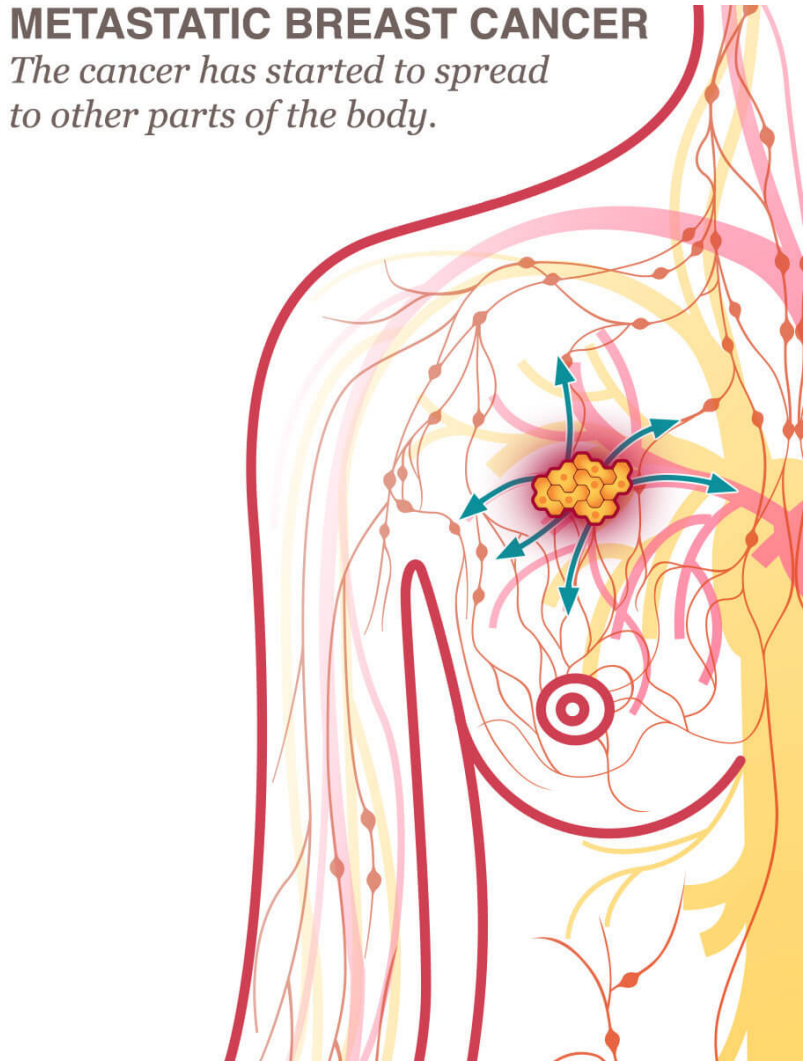
A tumour growth model that brings these components together can be fitted using maximum likelihood and any general-purpose optimiser (e.g. `optim` in R).

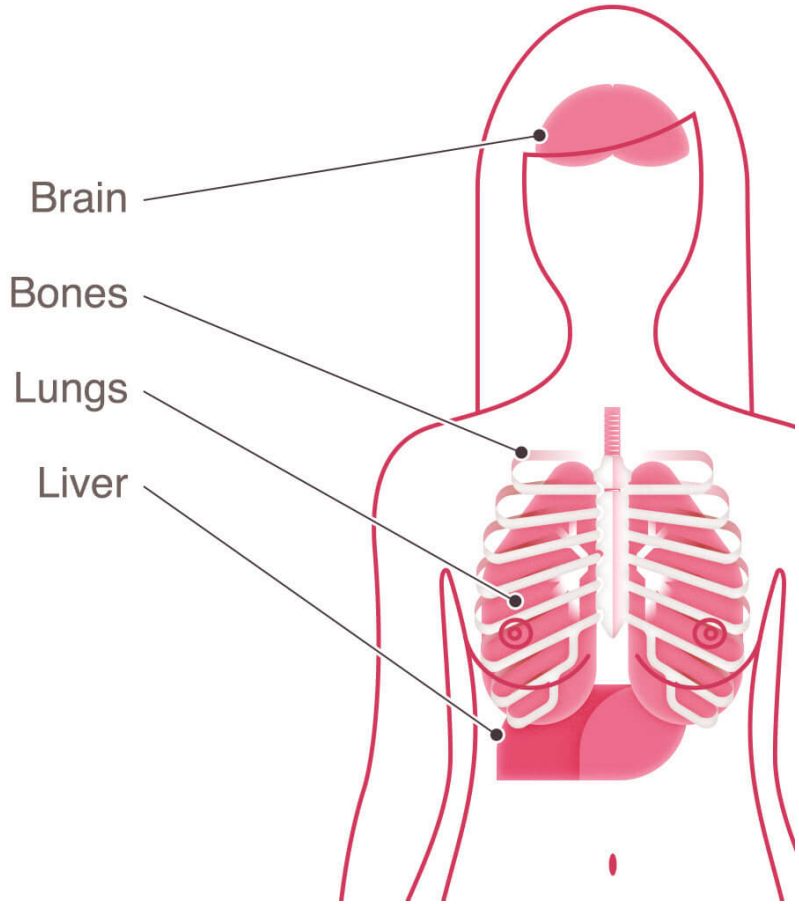
The likelihood function has closed-form, and it is based on:

1. The distribution of tumour sizes at screen detection
2. The distribution of tumour sizes at symptomatic detection

Both quantities are conditional on screening history; we skip the maths for simplicity, but more details can be found in Abrahamsson and Humphreys (2016) and Isheden and Humphreys (2019).

Breast Cancer Spread





Source: https://www.nationalbreastcancer.org

A Model for Successful Distant Metastatic Seeding

Metastatic seeding is assumed to follow an inhomogeneous Poisson process with intensity function

$$\lambda(t, r) = \sigma^* D(t, r)^k D'(t, r)$$

where $D(t, r)$ represents the number of times that cancer cells have divided, and $D'(t, r)$ represents the rate of cell division.

The exponent k (with $k \geq -1$) adds additional flexibility to the model.

We can identify a density and survival function based on this model, which conveniently follows a Weibull distribution.

A Model for Time to First Detected Distant Metastasis

Let W be the time from detection of the primary tumour to diagnosis of the first metastasis: $w = t + t_0 - t_{det}$.

Hazard and survival functions, alongside the density function of W, follow as:

$$S(w|V,R) = \exp\left[-\sigma\left(\frac{w}{r} + \log\frac{v_{det}}{V_0}\right)^{k+1}\right]$$

$$h(w|V,R) = \frac{\sigma}{r}(k+1)\left(\frac{w}{r} + \log\frac{v_{det}}{V_0}\right)^{k}$$

$$f_W(w|V,R) = h(w|V,R) \times S(w|V,R)$$

Model Assumptions

- It takes the individual-specific time t_{0i} for a metastasis to grow to a detectable size;
- Metastatic spread is independent of the detection probability of the primary tumour;
- Metastases growing to a detectable size before the diagnosis of the primary tumour can be visible at diagnosis, but not detected beforehand;
- Metastatic seeding stops when the primary tumour is diagnosed;
- Finally, we rely on *stable disease* assumptions.

Likelihood Function in the Absence of Screening

We can identify three types of observation that will contribute to the likelihood function:

1. Observed events, i.e. metastases that appear after diagnosis of the primary tumour ($w \geq 0$). These observations contribute $f_W(\cdot)$;
2. Left censored observations, i.e. individuals with detected metastases at the time of diagnosis ($w < 0$). These observations contribute the probability of having 1+ observed metastasis (derived from the inhomogeneous Poisson process);
3. Right censored observations, i.e. individuals that do not develop metastases before the end of follow-up. These observations will contribute the survival probability.

Preliminary Analysis

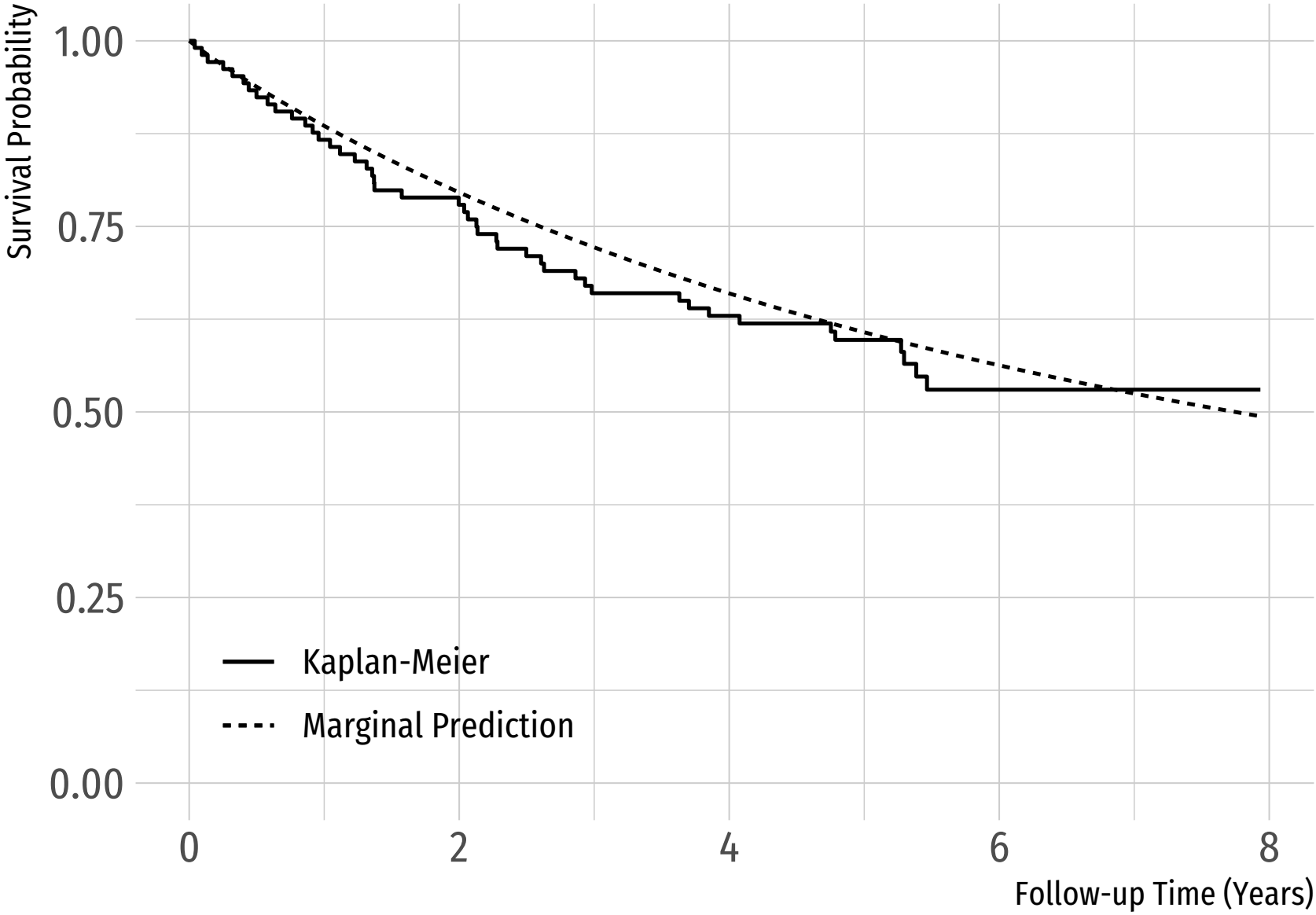
- We use data from CAHRES, a case-control study with recruitment in Sweden between 1993–1995;
- We extract 105 women with breast cancer detected symptomatically and with no screening history;
- Median tumour diameter was 25 mm (IQI: 17–35 mm);
- Median follow-up was 5.67 years (95% C.I. 5.24–5.94);
- 45 women (43%) showed distant metastasis before the end of the study.

Preliminary Analysis

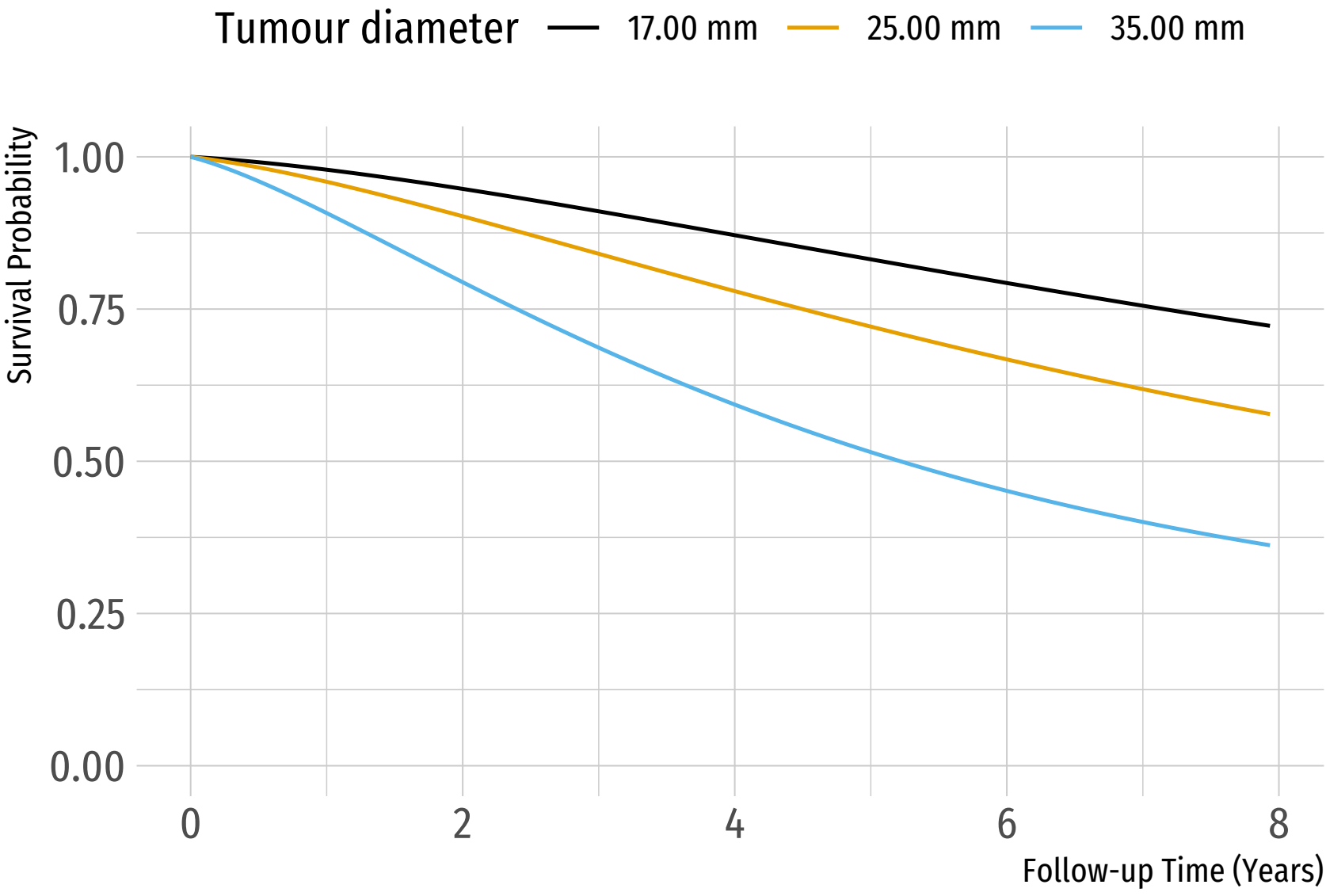
- We use data from CAHRES, a case-control study with recruitment in Sweden between 1993–1995;
- We extract 105 women with breast cancer detected symptomatically and with no screening history;
- Median tumour diameter was 25 mm (IQI: 17–35 mm);
- Median follow-up was 5.67 years (95% C.I. 5.24–5.94);
- 45 women (43%) showed distant metastasis before the end of the study.

Then, we fit the model described in the previous slides and we predict the probability of being free of distant metastasis over time.

Marginal Predictions



Conditional (On Tumour Size) Predictions



Coming Next: Extending the Model to Screening Data

The model we described so far is only valid in the absence of screening: we need to extend this to screening data.

Coming Next: Extending the Model to Screening Data

The model we described so far is only valid in the absence of screening: we need to extend this to screening data.

The likelihood function with screening data is formulated as a conditional joint probability:

$$P(\text{Size},\ \text{Metastasis} \mid \text{Mode of detection},\ \text{Screening History})$$

We need (again) stable disease assumptions, and we use a procedure similar to that described in Abrahamsson and Humphreys (2016) and Isheden and Humphreys (2019).

Conclusions

1. Continuous growth model can be used to model the unobserved natural history of a tumour. They are more efficient than Markov models when analysing several factors at once;
2. We introduce a model that relates the unobserved past natural history of tumours to the risk of future outcomes, e.g. distant metastatic spread. To the best of our knowledge, this is the first of its kind for breast cancer;
3. This is fundamentally important to assess new interventions and treatments e.g. in a microsimulation framework.