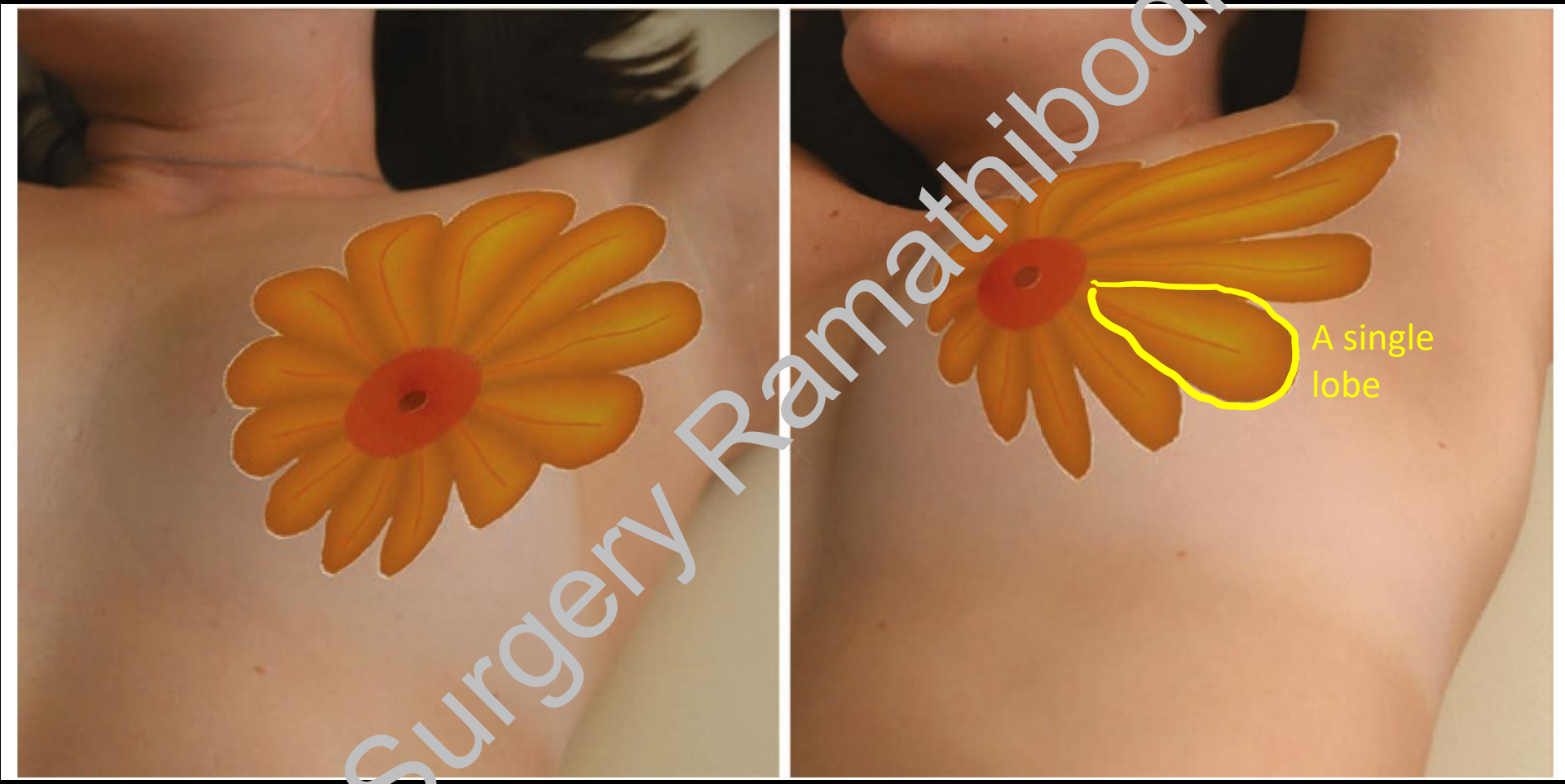
A Simple Mathematical Modeling of the Sick Lobe Hypothesis

Panuwat Lertsithichai

November 2018

Breast Lobes



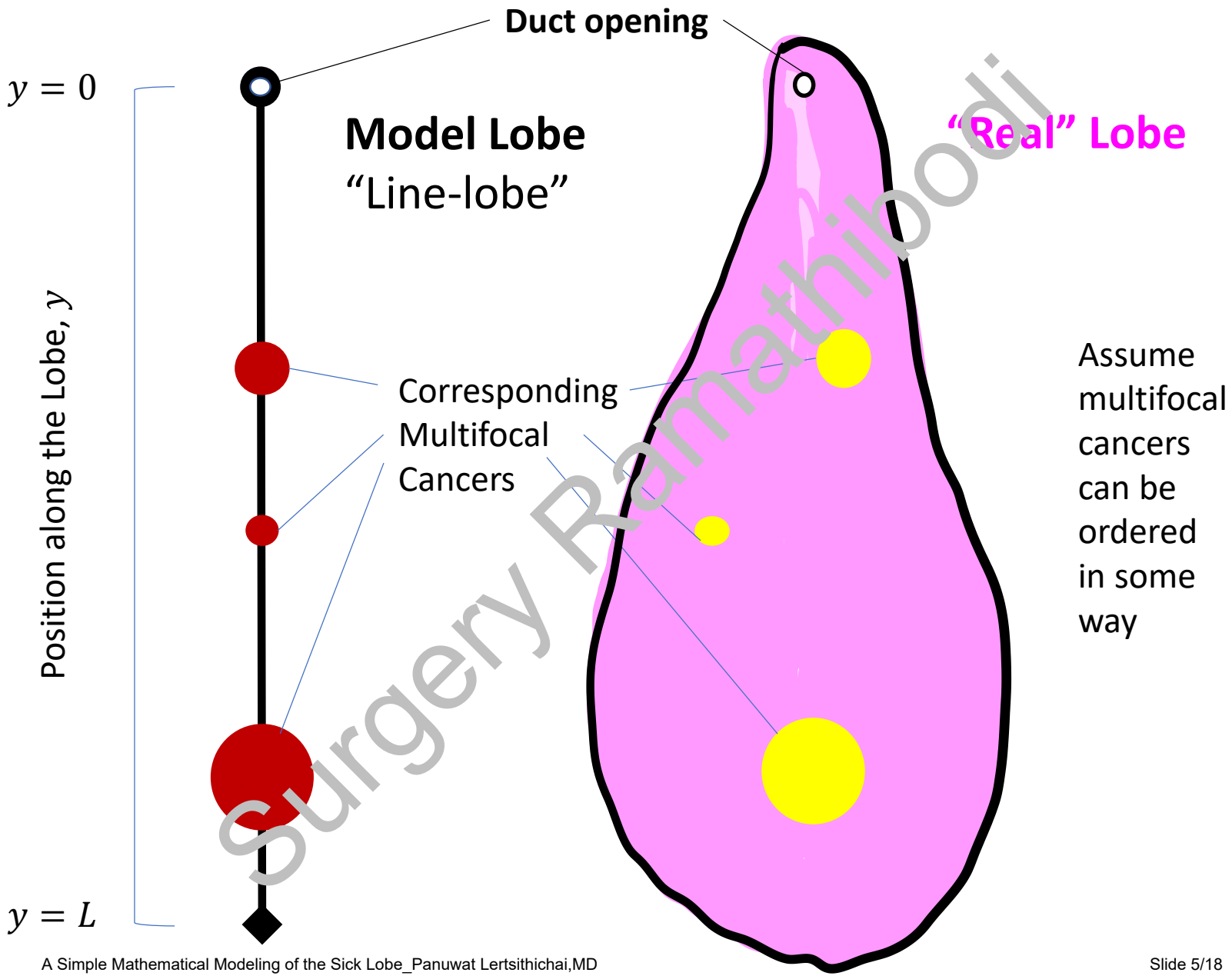
Dominique A, ed. Lobar Approach to Breast U/S. Springer, 2018

Sick Lobe Hypothesis

- Breast cancers usually occur in a breast lobe
- There is a genetic predisposition to cancer common to the whole lobe
- Thus **multifocal cancers** is common, but may sometimes be undetected, or is usually presented as a single focus of cancer

The Simplified Model

- **Linear model** (Dim=1) of the sick lobe ("line-lobe")
- Occurrence of cancer at any location/non-overlapping line segment is **independent** of one another (within the given lobe)
- Occurrences are **independent** in non-overlapping time intervals
- The *hazard of cancer* occurrence is **proportional** to length of line-lobe and follow-up time
- *Size of cancer* is a deterministic function of time



The Hazard

$$h(y,t) = 4\eta yt$$

- Where 4η is the *average number of cancer per unit length of the line-lobe, per unit time*
- The hazard *per lobe* at time t is thus

$$h(t) = 2\eta L^2 t$$

- Where L is the total length of the lobe
- Here, without loss of generality, we set $L = 1$

Average Number of Cancers

- Probability of having cancer within $[0, t]$ is
$$1 - e^{-\eta t^2}$$
- Expected number of cancers in a lobe found within $[0, t] \equiv E(N)_{[0,t]}$, is
$$\eta t^2$$
- Where the probability of $N = n$ follows the Poisson distribution with parameter ηt^2

Poisson Distribution

- Probability of Number of Cancers $N = n$ within any fixed time interval t follows a Poisson distribution if

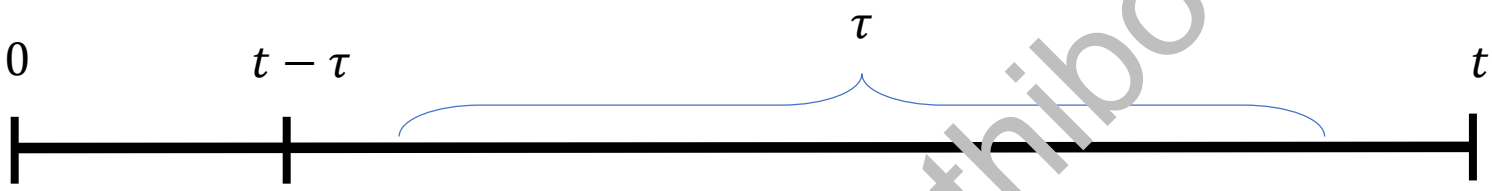
$$P(N = n) \equiv P(n) = e^{-\lambda_t} \frac{\lambda_t^n}{n!}$$

- Where the parameter $\lambda_t = \eta t^2$ is the average/ expected number of cancers occurring in time interval t
- The average of N is λ_t
- The variance of N is λ_t
- The probability of no cancer is $e^{-\lambda_t}$, etc.

Number of *Detected* Cancers

- Cancers are detected when they reach a certain size, say after a period τ
- Thus, cancers occurring for a period of time $\geq \tau$ will be (potentially) detectable
- Suppose a line-lobe has been followed for a time $t > \tau$, with a beginning at $time = 0$
- A patient with a sick lobe will have detectable cancers if at least one focus occurred within $[0, t - \tau]$

Number of Detected Cancers



Theorem*: the probability of n in $[0,t]$ (Poisson) can be written in terms of those in two independent, adjacent time segments (*non-overlapping partition*) $[0,t-\tau]$, $(t-\tau,t]$ as

- $p(n|0,t) = \sum_{j=0}^{n} \binom{n}{j} p(j|0,t-\tau)p(n-j|t-\tau,t)$

*Chapman-Kolmogorov

Number of Detected Cancers

Probability of n in $[0, t]$ with at least 1 cancer occurring in $[0, t-\tau]$:

- $p'(n|0,t) = p(n|0,t) - p(0|0,t-\tau)p(n|t-\tau,t)$

Expected number of cancers in $[0, t]$ with at least one in the period $[0, t-\tau] \equiv E(N|1_{[0,t-\tau]})_{[0,t]}$:

- $E(N|1_{[0,t-\tau]})_{[0,t]} = \sum_0^\infty p'(n|0,t)n / \sum_0^\infty p'(n|0,t)$

Where

$$p(0|0,t-\tau) = e^{-\eta(t-\tau)^2}$$

Number of Detected Cancers

Expected number of cancers in $[0, t]$ with at least 1 in the period $[0, t-\tau] \equiv E(N|1_{[0,t-\tau]})_{[0,t]}$ can be calculated as

$$E(N|1_{[0,t-\tau]})_{[0,t]} = \left(\eta t^2 - e^{-\eta(t-\tau)^2}\eta(t^2 - (t-\tau)^2)\right) / \left(1 - e^{-\eta(t-\tau)^2}\right)$$

- Which can be rewritten as:

$$\frac{\eta(t-\tau)^2}{\left(1 - e^{-\eta(t-\tau)^2}\right)} + \eta\left(t^2 - (t-\tau)^2\right)$$

- The first term is the expected number of cancers in $[0, t-\tau]$ given at least 1 had occurred by then;
- And the second is simply the expected number in the period $(t-\tau, t]$

Number of Detected Cancers

- In other words

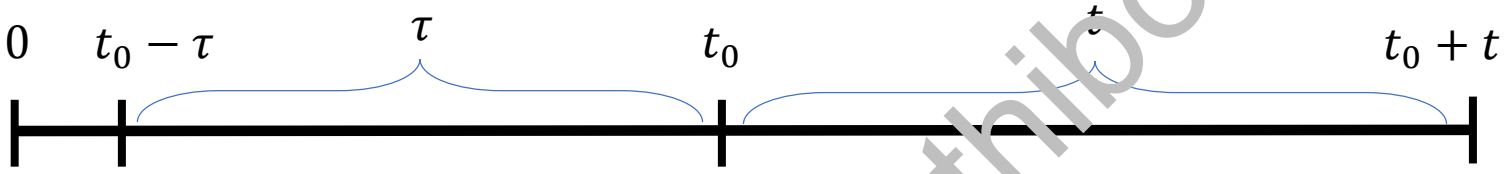
$$\frac{\eta(t-\tau)^2}{(1-e^{-\eta(t-\tau)^2})}$$

- Is the **expected number of detected cancers** in $[0, t]$ given at least one has occurred in $[0, t-\tau]$, and

$$\eta\left(t^2-(t-\tau)^2\right)$$

- Is the **expected number of undetected cancers** in the period $[0, t]$

Number of Multifocal Cancers & Background Risk



- A patient is found to have one detectable breast cancer at t_0 and is treated
- In the limit $\tau \to t_0$, expected number of cancers is $\eta {t_0}^2 + 1$ at $t_0 = \tau$
- Since only one of these is detected, the number of residual, multifocal cancers is $\eta \tau^2$
- This will be part of the **background risk** seen during the FU period, t

Simple Model of Residual Multifocal Cancer (Background Risk)

- If the "amount" of cancer per focus, e.g. **size of cancer**, is also required in addition to the number of residual foci when determining background risk
- The *simplest* but crudest way to incorporate this is to estimate the average size of a tumor occurring in an interval, then multiply by the expected number of cancers in that interval
- *A more complicated method* is to average a stochastic model

Average Amount of Residual Multifocal Cancer

Average amount occurring in a period $[0, t]$ conditional on at least one occurrence in that period, $E(I_{mJc}) =$

$$\sum_{n=1}^{\infty} \left\{ \int_{all\, t_i=0}^{t} \left(\sum_{i=1}^{n} s(t_i, t) \right) \boldsymbol{f}(t_i)\, dt_1 \dots dt_n \right\} p(n) / \sum_{n=1}^{\infty} p(n)$$

- Where $s(t_i, t)$ is the size of cancer at t given occurrence at $t_i < t$
- $p(n)$ is the probability of number of cancers n, given at least 1 in period $[0, t]$
- $\boldsymbol{f}(t_i) \equiv f(t_1, t_2, \dots, t_n)$ is the *multivariate probability density* of times of cancer occurrence $t_i < t$; i labels each of the n different times; all t_i's are independent; and the integration is n −dimensional

Average Amount of Residual Multifocal Cancer

Since all t_i's are independent,

- $\int_{all\,t_i=0}^{t}(\sum_{i=1}^{n} s(t_i,t))f(t_i)dt_1 \ldots dt_n = nE(S)$
- Where $E(S) = \int_0^t s(u,t)f(u)du$ is the expected size of a tumor occurring in $[0,t]$; thus
- $E(R_{mfc}) = \sum_{n=1}^{\infty} E(S)np(n) / \sum_{n=1}^{\infty} p(n)$; or
- $E(R_{mfc}) = E(S)E(N) = \frac{E(S)\eta t^2}{1-e^{-\eta t^2}}$; or, **if only one focus of cancer is detected and excluded:**
- $E(R_{mfc}) = E'(S)\eta\tau^2$; where $E'(S)$ excludes the detected cancer

Applications

- Need to fit parts of model to observational data first … (e.g. estimate values of γ & τ)
- The estimated amount of residual breast cancer after surgery (if significant) can be used to justify more aggressive *but* limited surgery (resection of the whole sick lobe)
- And thus to develop the technology to achieve true lobar resection
- Note that a more realistic model is also required!