Instantaneous geometric rates via generalized linear models

Andrea Discacciati   Matteo Bottai

Unit of Biostatistics
Karolinska Institutet
Stockholm, Sweden

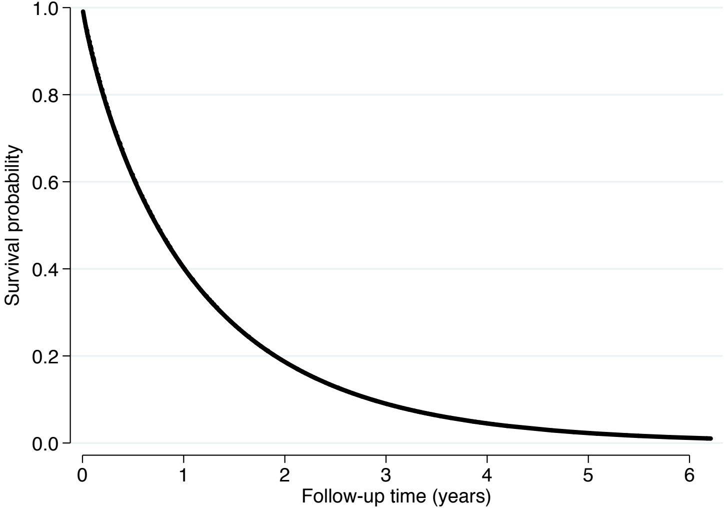
andrea.discacciati@ki.se

1 September 2017

Outline of this presentation

- Geometric rates
- Instantaneous geometric rates
- Models for the instantaneous geometric rates
- Instantaneous geometric rates via generalized linear models
- Example: survival in metastatic renal carcinoma
- Final remarks

Geometric rates

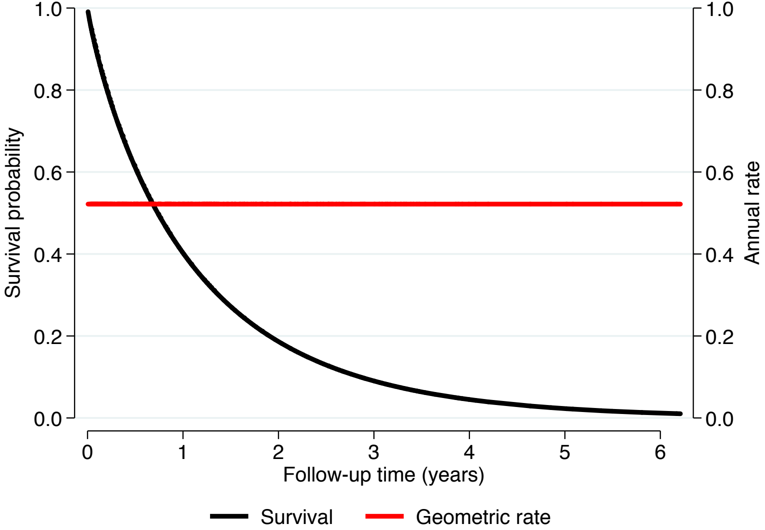


Geometric rates

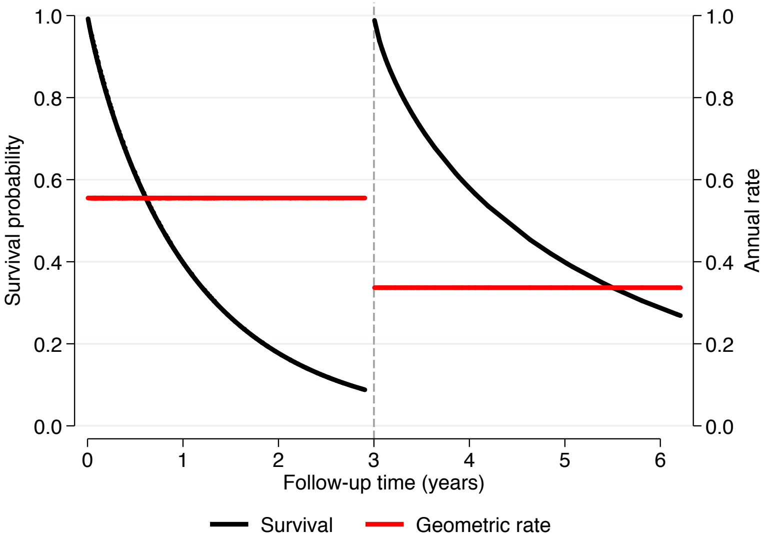
- The geometric rate represents the average probability of the event of interest per unit of time over a specific time interval $(0, t)$

$$g(0, t) = 1 - S(t)^{\frac{1}{t}}$$

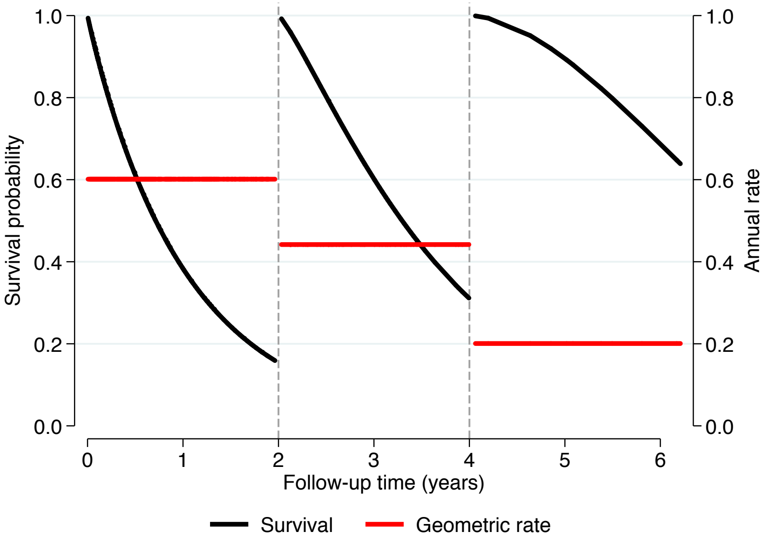
Geometric rates



Geometric rates



Geometric rates



Instantaneous geometric rates

- The instantaneous geometric rate (IGR) represents the instantaneous probability of the event of interest per unit of time

Instantaneous geometric rates

The limit of the geometric rate over shrinking time intervals $(t, t + \Delta t)$ gives the instantaneous geometric rate (Bottai, 2015)

$$\begin{aligned}
g(t) &\equiv \lim_{\Delta t \to 0^+} g(t, t + \Delta t) \\
&= \lim_{\Delta t \to 0^+} 1 - \left[\frac{S(t + \Delta t)}{S(t)} \right]^{\frac{1}{\Delta t}} \\
&= \lim_{\Delta t \to 0^+} 1 - \exp \left\{ \frac{\log S(t + \Delta t) - \log S(t)}{\Delta t} \right\} \\
&= 1 - \exp \left\{ \frac{\partial \log S(t)}{\partial t} \right\} \\
&= 1 - \exp \left\{ -\frac{f(t)}{S(t)} \right\} \\
&= 1 - \exp \{ -h(t) \}
\end{aligned} \tag{1}$$

Models for the instantaneous geometric rate

- Proportional instantaneous geometric rate model

$$g_i(t|\boldsymbol{x}_i) = g_0(t)\exp\{\boldsymbol{x}_i^T\boldsymbol{\beta}\} \tag{2}$$

- Proportional instantaneous geometric odds model

$$\frac{g_i(t|\boldsymbol{x}_i)}{1-g_i(t|\boldsymbol{x}_i)} = \frac{g_0(t)}{1-g_0(t)}\exp\{\boldsymbol{x}_i^T\boldsymbol{\beta}\} \tag{3}$$

- These models can be estimated within the generalized linear model (GLM) framework by using two nonstandard link functions

Instantaneous geometric rates via GLM

Let's focus on the proportional instantaneous geometric rate model

By taking the logarithm of both sides of (2) we get

$$\log[g_i(t|\boldsymbol{x}_i)] = \log[g_0(t)] + \boldsymbol{x}_i^T\boldsymbol{\beta}$$

and by equation (1) we write

$$\log[1 - \exp\{-h_i(t)\}|\boldsymbol{x}_i] = \log[g_0(t)] + \boldsymbol{x}_i^T\boldsymbol{\beta} \tag{4}$$

where $\log[g_0(t)]$ (baseline log-IGR) is modelled using for example polynomials or splines.

Instantaneous geometric rates via GLM

- To model the baseline log-IGR, we split each individual's follow-up time into very short intervals (`stsplit`)
- Given equation (4) we use the following link function

$$\eta_{ij} \equiv k(\mu_{ij}) = \log\left[1 - \exp\left\{-\frac{\mu_{ij}}{t_{ij}}\right\}\right]$$

where:

- t_{ij} is the length of the jth interval relative to the ith subject
- μ_{ij} is the expected value of d_{ij} (the event/censoring indicator), which is assumed to follow a distribution of the exponential family

Instantaneous geometric rates via GLM

- In model (2) the exponentiated coefficients $\exp\{\beta\}$ are interpreted as instantaneous geometric rate ratios, whereas in model (3) they are interpreted as instantaneous geometric odds ratios
- If the instantaneous geometric rates are proportional across different populations, the instantaneous geometric odds are not, and vice-versa

- Link functions for models (2) and (3) are implemented in two user-defined link programs: `log_igr` and `logit_igr`

Example: survival in metastatic renal carcinoma

- Data from a clinical trial on 347 patients diagnosed with metastatic renal carcinoma
- The patients were randomly assigned to either interferon-α (IFN) or oral medroxyprogesterone (MPA)
- A total of 322 patients died during follow-up

- Stata code to reproduce the worked-out example is available at:

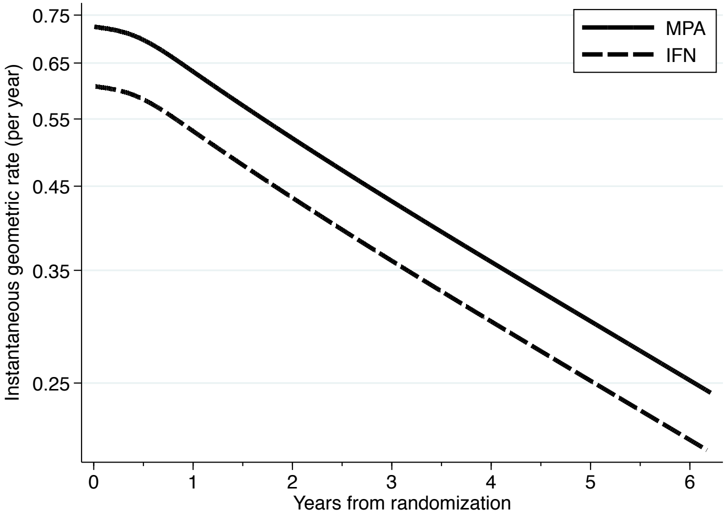
  `www.imm.ki.se/biostatistics`

Example: survival in metastatic renal carcinoma

```
. qui use http://www.imm.ki.se/biostatistics/data/kidney, clear
. qui stset survtime, failure(cens) id(pid) scale(365.24)
. qui stsplit click, every(`=1/52')
. qui generate risktime = _t - _t0
. qui rcsgen _t, df(3) if2(_d == 1) gen(_rcs)

. glm _d i.trt c._rcs?, family(poisson) link(log_igr risktime) vce(robust) ///
>         nolog search eform baselevel noheader

initial:       log pseudolikelihood =     -<inf>  (could not be evaluated)
feasible:      log pseudolikelihood = -4804.4455
rescale:       log pseudolikelihood = -1959.6083

------------------------------------------------------------------------------
             |               Robust
          _d |     exp(b)   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
         trt |
        MPA  |       1.00  (base)
        IFN  |       0.84       0.06    -2.62   0.009          0.73        0.96
             |
       _rcs1 |       0.96       0.29    -0.13   0.894          0.53        1.74
       _rcs2 |       1.31       0.86     0.41   0.681          0.36        4.72
       _rcs3 |       0.90       0.24    -0.39   0.693          0.54        1.51
       _cons |       0.72       0.06    -3.63   0.000          0.61        0.86
------------------------------------------------------------------------------
```

- IGRR comparing IFN versus MPA patients: 0.84 (0.73–0.96)

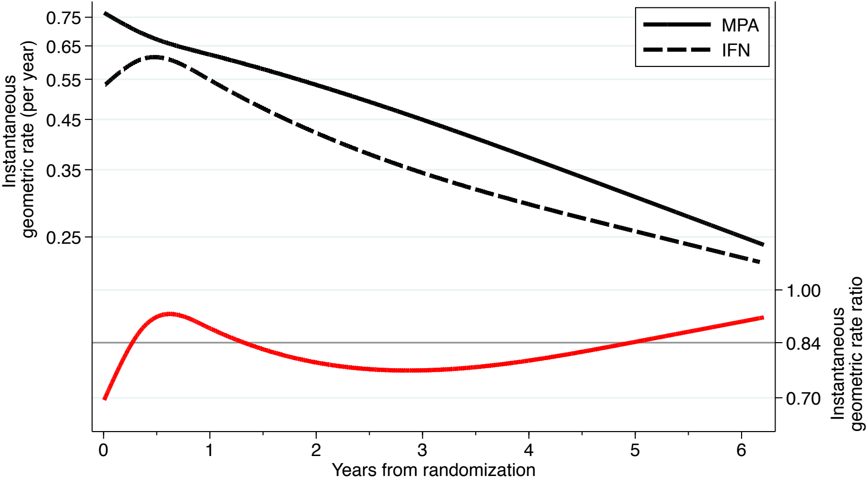
Example: survival in metastatic renal carcinoma



Example: survival in metastatic renal carcinoma

```
. glm _d i.trt##c._rcs?, family(poisson) link(log_igr risktime) vce(robust) ///
>         nolog search baselevel noheader

initial:       log pseudolikelihood =     -<inf>  (could not be evaluated)
feasible:      log pseudolikelihood = -4804.4455
rescale:       log pseudolikelihood = -1959.6083

------------------------------------------------------------------------------
             |               Robust
          _d |      Coef.   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
         trt |
        MPA  |       0.00  (base)
        IFN  |      -0.37        0.19    -1.92   0.054        -0.74        0.01
             |
       _rcs1 |      -0.31        0.36    -0.86   0.389        -1.02        0.40
       _rcs2 |      -0.29        0.82    -0.36   0.722        -1.91        1.32
       _rcs3 |       0.12        0.33     0.35   0.727        -0.54        0.77
             |
 trt#c._rcs1 |
        IFN  |       0.78        0.66     1.18   0.238        -0.52        2.08
             |
 trt#c._rcs2 |
        IFN  |       1.57        1.38     1.14   0.256        -1.14        4.29
             |
 trt#c._rcs3 |
        IFN  |      -0.62        0.55    -1.11   0.267        -1.70        0.47
             |
       _cons |      -0.26        0.09    -3.02   0.003        -0.44       -0.09
------------------------------------------------------------------------------
```

Example: survival in metastatic renal carcinoma



Final remarks

- Instantaneous geometric rates are easy to interpret
- Instantaneous geometric rates $\neq$ hazard rates

- Proportional instantaneous geometric rate/odds models for the effect of covariates on the IGR
- These models can be estimated within the GLM framework by using nonstandard link functions

References

- Bottai M. A regression method for modelling geometric rates. *Stat Methods Med Res*. 2015 Sep 18.
- Discacciati A, Bottai M. Instantaneous geometric rates via generalized linear models. *Stata J*. 2017;17(2):358–371.