Self-controlling case series analysis

January 2007

Niels Keiding

Self-controlling case series analysis

C.P. Farrington & H.J. Whitaker (2006). *Appl. Statist.* **55**, 553-594 (with discussion).

Recurrent events for individual *i* in interval $\left(a_i, b_i\right]$ occur with intensity $\lambda_i\left(t \mid x_i^t\right)$ where $x_i^t = \left\{x_i\left(s\right) : s \le t\right\}$ is the exposure history at *t*. Assume $\left(x_i(s)\right)$ *exogenous*, i.e.

$$\lambda_i\left(t \mid x_i^t\right) = \lambda_i\left(t \mid x_i\right) \text{ where } x_i = x_i^{\infty}$$

Then given x_i, the counting process for individual *i* is a non-homogeneous Poisson process with intensity $\lambda_i\left(t \mid x_i\right)$ and the likelihood for experiencing events at times $t_{i1}, \cdots, t_{in_i}$ is

$$\prod_{j=1}^{n_i} \lambda_i\left(t_{ij} \mid x_i\right) \exp\left\{-\int_{a_i}^{b_i} \lambda_i\left(t \mid x_i\right) dt\right\}$$

Conditional likelihood, proportional incidence model

Consider likelihood conditional on number n_i of events:

$$L_i^c = \prod_{j=1}^{n_i} \lambda_i\left(t_{ij} \mid x_i\right) \Big/ \left\{\int_{a_i}^{b_i} \lambda_i\left(t \mid x_i\right) dt\right\}^{n_i}.$$

Note $L_i^c = 1$ when $n_i = 0$: case-only!
Assume *proportional incidence:*

$$\lambda_i\left(t \mid x_i\right) = \varphi\psi(t)\exp\left(\gamma_i + x_i(t)\beta\right)$$

with underlying incidence φ, relative incidence $\psi(t)$, γ_i =random and fixed individual effects that are constant over *t*. Parameter of interest: β. With N cases we get

$$L = \prod_{i=1}^{N}\prod_{j=1}^{n_i} \left\{\psi\left(t_{ij}\right)\exp\left(x_i\left(t_{ij}\right)\beta\right) \Big/ \int_{a_i}^{b_i} \exp\left(x_i\left(t_{ij}\right)\beta\right)\psi\left(t\right)dt\right\}$$

so φ and γ_i have disappeared. Easiest to assume ψ completely arbitrary (*semi-parametric model*).

Assumptions

1. Exogeneity $\lambda_i\left(t \mid x_i^t\right) = \lambda_i\left(t \mid x_i\right)$

2. Recurrent events or *rare* non-recurrent events (see later)

3. Multiplicative intensity (otherwise time-independent covariates do not cancel out)

4. Observation periods $\left(a_i, b_i\right]$ independent of event times. (Stronger than independent censoring)

Example: side effects of MMR vaccine

Idiopathic thrombocytopenic purpura (ITP) is an uncommon, potentially recurrent bleeding disorder. Effect (if present) is transient. But not considered contraindication to vaccination, so there should be no feedback (no endogeneity).

35 children admitted to hospital with ITP during age 366-730 days.

29: one event, 5: two events, 1: five events. Possibly different baseline incidence – but this is self-controlled. All had had MMR vaccination

$$x_i(t) = \begin{cases} 2 & \text{if case } \in [0,42] \text{ days after vaccination ("exposed")} \\ 1 & \text{otherwise ("unexposed")} \end{cases}$$

There were 31 unexposed, 13 exposed.

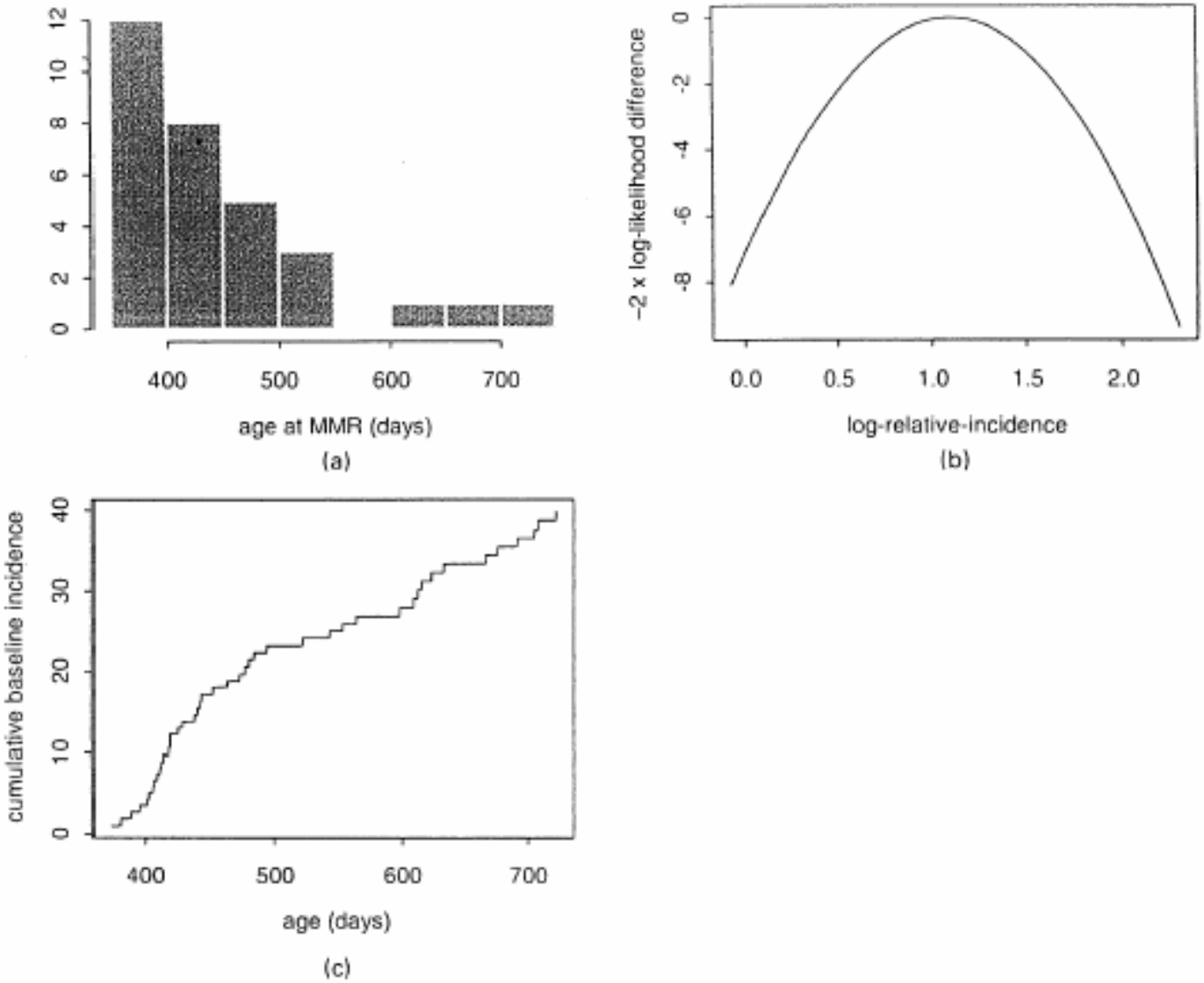

Fig. 1. ITP and MMR vaccination: (a) histogram of age at MMR vaccination in exposed cases; (b) profile log-likelihood for log-relative-incidence; (c) estimated cumulative relative base-line incidence

Example: side effects of MMR vaccine

Vaccine effect $\exp\left(\hat{\beta}\right) = 3.01 \quad (1.38, 6.54)$

Profile likelihood ratio test for $\exp\left(\beta\right) = 1$: $P = 0.008$

Independent nested case-control study of first ITP: $\exp\left(\hat{\beta}\right) = 6.3 (1.3, 30.1)$:

more narrow confidence interval on case-only study: there were many "exposed" individuals and risk period short relative to observation period.

Case series model for rare non-recurrent events

Now $\lambda_i\left(t \mid x_i^t\right) = \lambda_i\left(t \mid x_i\right)$ is the hazard function with $x_i = x_i^{b_i}$ with b_i large, let $S_i\left(t \mid x_i\right)$ denote the survival function. Given x_i and given that an event occurred in $\left(a_i, b_i\right]$, the likelihood for i is

$$L_i = S_i\left(t_i \mid x_i\right)\lambda_i\left(t_i \mid x_i\right) \Big/ \left\{S_i\left(a_i \mid x_i\right) - S_i\left(b_i \mid x_i\right)\right\}.$$

Assume $\lambda_i\left(t \mid x_i\right) = \varphi\nu_i\left(t \mid x_i\right)$ with the relative hazard ν_i bounded on $\left(0, b_i\right]$.

As $\varphi \to 0$, $S_i\left(t \mid x_i\right) \to 1$ and

$$S_i\left(a_i \mid x_i\right) - S_i\left(b_i \mid x_i\right) \to \int_{a_i}^{b_i} \lambda_i\left(t \mid x_i\right) dt$$

so L_i tends to the likelihood studied earlier.

Example: MMR vaccine and autism

Suspicion raised in 1998.
357 cases of autism diagnosed in children up to 16 years of age.

Table 5. Relative incidence of autism after first MMR vaccine

Risk period (months)	*Events*	*Relative incidence (95% confidence interval)*
Unexposed	94	1 (—)
0–24 post MMR vaccination	131	0.892 (0.400, 1.99)
25–48 post MMR vaccination	109	0.755 (0.310, 1.84)
49–72 post MMR vaccination	17	0.849 (0.275, 2.63)
⩾72 post MMR vaccination	6	0.903 (0.236, 3.45)

Combined post-MMR estimate 0.882 (0.399, 1.95)

Table 6. Three studies of autism and MMR

Study	*Sample size*	$exp(\hat{\beta})$	*95% confidence interval*
Case series	357 cases	0.88	(0.40, 1.95)
Cohort	537303 children, 316 cases	0.92	(0.68, 1.24)
Case–control	1294 cases, 4469 controls	0.86	(0.68, 1.09)

Cohort study: Madsen et al. (2002) SSI.