Finite Mixture Models to Identify Reverse Causation CT Scans

Nicolas Roydon Smoll, John D. Mathews, Katrina J. Scurrah

University of Melbourne, School of Population and Global Health

Introduction

Cohort studies of large populations exposed to computed tomography (CT) scans in childhood or adolescence have shown increased rates of leukemia, brain cancer, and most other cancers.[1, 2]

- CT scans undertaken shortly before cancer diagnosis are usually part of that diagnostic process (*reverse causation (RC) exposures*)
- Exclusion and lag periods are used to reduce the effect of RC CT scan exposures (Exclusion means person has zero dose).
- Finite mixture models are used to provide an evidence-based exclusion or lag period

Finite Mixture Model

Finite mixture models (FMMs) are used to model the probability of an observation belonging to an unobserved variable, termed a latent class, such as honest reporting of drug use, medical literacy, the likelihood of seeking help at the first sign of an illness, or as in this study, the reason a CT scan was performed (diagnostic process vs. other reason).

Methods

- De-identified records of CT scan exposures for Australians aged 0–19 years and registered with Medicare (Australia's universal healthcare system)
- Generalized linear model with an exponential distribution for *interval times* (time between CT scan and diagnosis of brain cancer)

Classifier Methods

1. Crossover method: The classifier is the interval time at which the posterior probability of being classified in the late class is greater than the posterior probability of being classified in the earlier, or RC class.
2. 99th Percentile Method: The classifier is the interval time at which the posterior probability of being classified in the late class is greater than the 99th percentile.

Results

- Australian Medicare cohort is comprised of 11,528,078 persons.
- 1028 are children with a diagnosis of a brain tumor and have a recorded CT scan before the diagnosis.
- Total of 1450 recorded CT scans, and 1255 (87%) were CT scans of the brain.

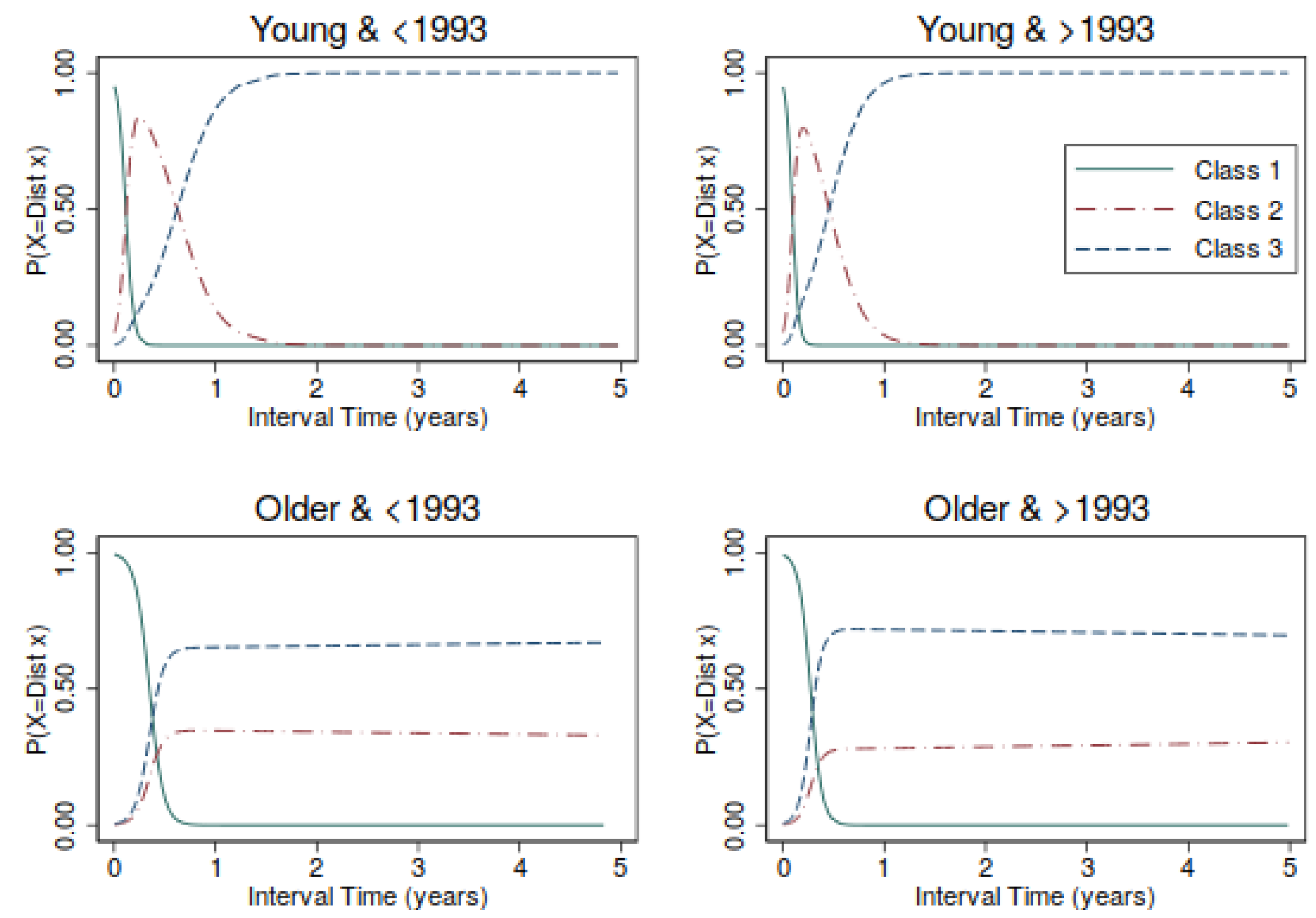

Fig. 1: FMM with three latent classes, later collapsed to 2 classes: reverse causation and potentially causal class. In the young, the first 2 classes were collapsed and in the older group, the second two classes were collapsed. Modelled alone, the older group found only two classes.

Classifiers

The crossover classification method classifiers argue for an exclusion period greater than or equal to 8.7 months. **See where the lines cross in figure 2**.
The 99th percentile classifiers argue for an exclusion period greater than or equal to 18.9 months. **See where the early distribution line approximates a probability of $\leq$ 0.01 in figure 2**.

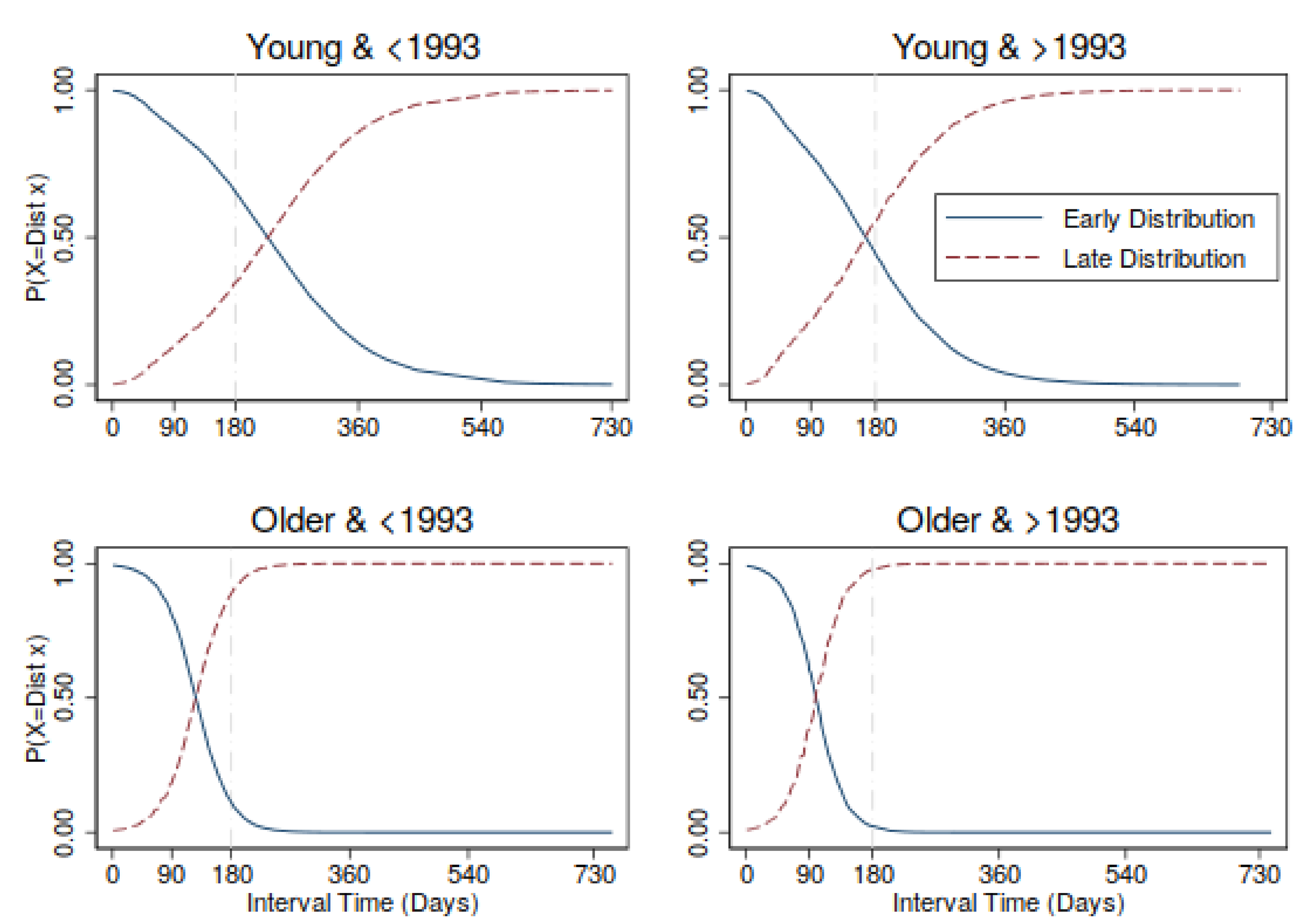

Fig. 2: Posterior Probabilities. The posterior probability of being in the first two classes (RC class) decreases to less than 1 percent at 19 months (young group before 1993). This cutoff is as short as 6.8 months in the older group after 1993.

Discussion

The model suggests that CT scans occurring after one or two years are *potentially causal* CT scans. The results of the external validation suggest that the 99th percentile classifier is optimal and exclusion periods of two years should be used.

External Validation

A systematic review was of the pre-diagnostic symptomatic interval (PSI; time between first symptom and diagnosis) for children with brain tumors. The goal was to calculate the sample-size weighted mean of the 99th percentile of the PSI for children with brain tumors and use this to inform our choice of classification method. The sample-size weighted mean of the P_{99} was 15.6 months. This supports the findings of the 99th percentile classifier.

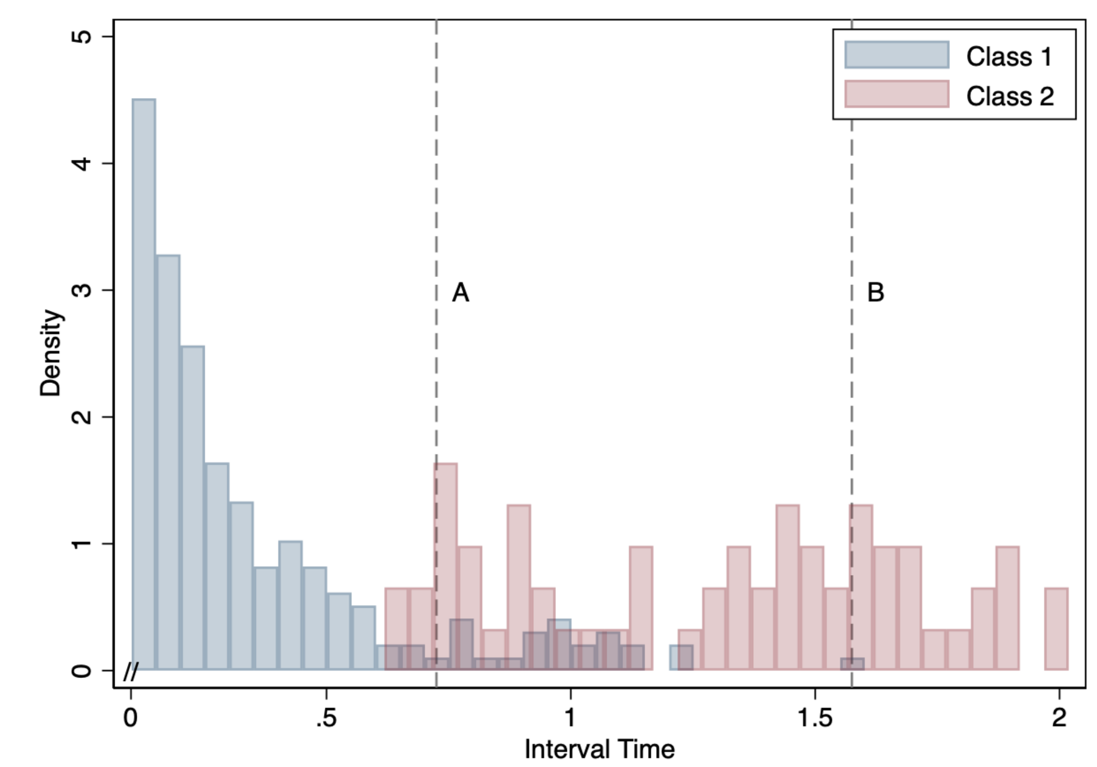

Fig. 3: Distribution after classification. A = suggested crossover classifier, B = suggested P_{99} classifier, cæsura represents where first 5 weeks of scans deleted to better view remaining distribution. Remember the precise cutpoint depends on age and year of CT scan.

Acknowledgements

The Australian government provided de-identified Medicare records; state and territory governments provided cancer and death records and approved linkage by the Australian Institute of Health and Welfare. The research was funded by the Australian Government through NHMRC grants 509190, 1027368, and 1084197. We thank Anna Forsythe, Zoe Brady and the many other colleagues who have made this report possible.

References

[1] John D Mathews et al. "Cancer risk in 680 000 people exposed to computed tomography scans in childhood or adolescence : data linkage study of 11 million Australians". In: *British Medical Journal* (2013), pp. 1–18.

[2] NR Smoll, K Scurrah, and JD Mathews. "Exposure to ionizing radiation and brain cancer incidence: The Life Span Study cohort". In: *Cancer Epidemiology* 42 (2016), pp. 60–65.