

Statistical Methods in Epidemiology

Lab 6

Case–Control Studies

Στόχος:

Να κατανοηθούν τα βασικά στοιχεία σχεδιασμού και ανάλυσης case–control μελετών, με έμφαση: α) στη λογική του matched case–control, β) στις τεχνικές conditional logistic regression, γ) στον εντοπισμό συγχύσεως και τροποποίησης επίδρασης, δ) στις μεθοδολογικές επιλογές για την επιλογή των controls (cumulative sampling, case–cohort, incidence density sampling).

I. Matched Case–Control: Salmonella Typhimurium Dataset

Το φθινόπωρο του 1996, ένας ασυνήθιστα μεγάλος αριθμός κρουσμάτων Salmonella Typhimurium καταγράφηκε στη Φινλάνδία (Fyn) της Δανίας. Το Danish Zoonosis Centre οργάνωσε μια matched case–control μελέτη, προκειμένου να εντοπίσει τις πηγές. Για κάθε case βρέθηκαν δύο controls εξομοιωμένα ως προς την ηλικία, το φύλο και τον τόπο κατοικίας. Οι συμμετέχοντες απάντησαν σε τηλεφωνική συνέντευξη σχετικά με την κατανάλωση τροφίμων κατά τις τελευταίες δύο εβδομάδες. Τα δεδομένα από αυτή τη μελέτη βρίσκονται στο αρχείο δεδομένων typhi.dta.

1. Ανοίξτε το αρχείο **typhi.dta**. Ποιες μεταβλητές περιλαμβάνει;

```
. use typhi, clear  
. desc
```

2. Εξετάστε την επίδραση του pork με conditional logistic regression. Τι δείχνει η εκτίμηση;

```
. clogit case pork, group(set)
```

3. Ελέγξτε όλες τις υπόλοιπες μεταβλητές τροφίμων. Ποιες μεταβλητές είναι στατιστικά σημαντικές;

```
. clogit case beef, group(set) nolog  
. clogit case poultry, group(set) nolog  
. clogit case liverp, group(set) nolog  
. clogit case meat, group(set) nolog  
. clogit case veg, group(set) nolog  
. clogit case fruit, group(set) nolog  
. clogit case egg, group(set) nolog  
. clogit case plant7, group(set) nolog
```

4. Εξετάστε εάν υπάρχει αλληλεπίδραση μεταξύ του fruit (προστατευτικός παράγοντας) και plant7 (παράγοντας κινδύνου).

α) Εκτιμήστε μοντέλο με τους δύο κύριους όρους.

```
. clogit case fruit plant7, group(set) nolog
```

β) Εκτιμήστε μοντέλο με τον όρο αλληλεπίδρασης.

```
. xi: clogit case i.fruit*i.plant7, group(set) nolog  
. char plant7 [omit] 1  
. xi: clogit case i.fruit*i.plant7, group(set) nolog
```

γ) Συμπληρώστε τον παρακάτω πίνακα των log(OR)

log(OR)	fruit = 0	fruit = 1
plant7 = 0		
plant7 = 1		

δ) Ερμηνεύστε τους συντελεστές.

5. Διερευνήστε την υπόθεση ότι “το fruit είναι προστατευτικό μόνο μεταξύ όσων έχουν καταναλώσει meat από plant7”.

α) Πρόκειται για αλληλεπίδραση;

β) Πώς θα παραμετροποιούσατε ένα τέτοιο μοντέλο ώστε η ομάδα αναφοράς να είναι: plant7 = 1 & fruit = 0;

log(OR)	fruit = 0	fruit = 1
plant7 = 0		
plant7 = 1		

6. Δημιουργήστε μεταβλητή με τρεις κατηγορίες που συνδυάζει fruit και plant7, και συγκρίνετε την επίδραση των κατηγοριών. Ερμηνεύστε συγκριτικά με το ερώτημα 4.

```
. gen frpl1 = 0 if (fruit == 0 & plant7 == 0) | (fruit == 1 & plant7 == 0)
. replace frpl1 = 1 if (fruit == 0 & plant7 == 1)
. replace frpl1 = 2 if (fruit == 1 & plant7 == 1)

. xi: clogit case i.frpl1, group(set) nolog
```

II. Επιλογή Controls σε Case–Control Μελέτες

Στην ενότητα αυτή θα αναλύσετε αποσπάσματα από τις μελέτες:

1. **Mahmood et al. (1989)** – infant feeding & diarrhea
2. **Mueller et al. (1987)** – tonsillectomy & Hodgkin’s disease
3. **Collins et al. (1999)** – surgery & Creutzfeldt–Jakob disease (CJD)

και θα απαντήσετε στα ερωτήματα που ακολουθούν, με βάση την ταξινόμηση των sampling schemes:

- Cumulative (exclusive) sampling
- Case–cohort sampling
- Incidence density (risk-set) sampling

1. Μελέτη Mahmood et al. (1989)

- (α) Από ποιον πληθυσμό στρατολογήθηκαν οι controls;
- (β) Ποια είναι τα πλεονεκτήματα και μειονεκτήματα αυτού του τρόπου επιλογής;
- (γ) Ποιο sampling scheme χρησιμοποιήθηκε; Τεκμηριώστε.
- (δ) Ποιο μέτρο σχετικής επίπτωσης εκτιμάται (risk ratio, rate ratio ή odds ratio); Γιατί;
- (ε) Γιατί αποκλείστηκαν τα παιδιά ≥ 3 μηνών χωρίς ιστορικό επίσκεψης για εμβολιασμό;

2. Μελέτη Mueller et al. (1987) – Χρήση siblings ως controls

- (α) Ποια είναι τα πλεονεκτήματα αυτής της επιλογής;
- (β) Ποια είναι τα μειονεκτήματα; Πώς οδηγούν σε πιθανή μεροληψία;

3. Μελέτη Collins et al. (1999) – Random telephone sampling

- (α) Γιατί προτιμήθηκε αυτή η προσέγγιση αντί για νοσοκομειακούς controls;
- (β) Ποιες μορφές μεροληψίας μπορεί να εισάγει το random telephone sampling;

III. Nested Case–Control & Case–Cohort Σχεδιασμοί

Δεδομένης της θεωρίας (SME_6_nested_CC):

1. Περιγράψτε συνοπτικά τη διαφορά μεταξύ nested case–control και case–cohort.
2. Σε ποια περίπτωση το incidence density sampling δίνει OR που προσεγγίζει το rate ratio;
3. Πότε το cumulative sampling επιτρέπει την εκτίμηση του risk ratio;
4. Ποια είναι τα πλεονεκτήματα χρήσης nested case–control έναντι πλήρους cohort ανάλυσης;