Exercise 3.1 Inclusive, exclusive and concurrent sampling in Excel

The Excel sheet provided contains an ideal population of 10,000 exposed and 10,000 unexposed individuals followed over a 10-year period, where each year the incidence of a disease was exactly 5% and 1% of the exposed (pink) and unexposed(blue) persons at risk. These rates were used in the equations that generated column B and column E.

Check that you understand cells B16-B18 which calculate the cumulative incidence proportion, cumulative odds and the incidence (per person year) where it has been assumed that all individuals in the mid-year population contributed one full year. From these three quantities, the RR, cumulative incidence OR and the IRR in the small highlighted box are calculated.

Check that you understand how the odds ratios from inclusive, exclusive and concurrent sampling are calculated in the larger highlighted box (as you can see, these three ORs correspond to the RR, cumulative incidence OR and IRR).

Now alter the incidence in the exposed and unexposed to some other values, by changing the values in cells B3 and E3 and copying these cells down to row 12.

Verify that

    (i)   the OR from concurrent sampling provides a good estimate of your IRR (i.e. the ratio of the incidences that you used for the exposed and unexposed).

    (ii)   the OR from exclusive sampling overestimates the IRR