# Computational Epidemiology
## Assignment #4

You must submit your **programs** (spreadsheets, python programs, or Java programs) and **report.pdf** containing a short report for each of the question. You should clearly describe your approaches that have led to your findings.

1. Suppose a random walker moves in a grid of size 100 x 100 cell. The walker will start position 50 x 50 and will make 1000 moves. Randomly distributed in the grid are n (n= 1, 5, 50) contaminants or sources of infectious material. Whenever the random walker reaches a grid position with such a contaminant it is marked as infected and will stay infected for the remainder of the walk (so you could actually terminate the walk at this point). Conduct computational experiments to explore the following:

    a) The probability of the walker to become infected as a function of the number of contaminants in the grid.
    b) The probability of the walker to become infected as a function of the length (i.e., number of steps) of the random walk.
    c) The probability of the walker to become infested as a function of "average step-size". For this you need to alter the average distance covered by each step the walker makes. The actual step size of each step must be an integer and is chosen from a distribution.

Use any spreadsheet application to graph the change in probability of becoming infected. You must briefly describe your findings. How do you address the issues of a finite grid, i.e., the walker could "fall-off?"

2. As an epidemiologist you are going to investigate the effect of a drug suspected of causing malformations in newborn infants when the drug in question is taken by pregnant women during the course of their pregnancy. As your sample you will use the next 200 single births occurring in a given hospital. For each birth a medication history will be taken from the new mother and from her doctor. [N.B.: These mothers are considered to have been followed prospectively during the entire course of their pregnancy, because a complete and accurate record of drug use was maintained during pregnancy.]

    The resultant data are:
    Forty (40) mothers have taken the suspected drug during their pregnancy.
    Of these mothers, 35 have delivered malformed infants. In addition, there are 10 other infants born with malfunctions or malformations.
    a) What is the number of mothers who did not take the drug and did not deliver a malformed infant?
    b) What is the exposure variable of this study?
    c) What is the outcome variable of this study?
    d) What is the incidence in the exposed?
    e) What is the incidence in the non-exposed?
    f) What measure of association are you going to use in this study?
    g) Use a 2 x 2 table and interpret the results of calculating the measure of association.

3. A case-control study was conducted to evaluate the association between oral contraceptive use and breast cancer risk. About 215 breast cancer cases were identified from the city cancer registry and 690 controls were selected from the city population. The use of oral contraceptives was recalled

(remembered) by 140 cases and 436 controls. Set up the 2 x 2 Table and calculate the measure of association, provide interpretation.

4. Which type of study is being described in each of the following scenarios (CASE-CONTROL, RETROSPECTIVE COHORT, ECOLOGIC, OR CROSS-SECTIONAL)? **Justify your answer**.
   a. A study that examines the death rates from cervical cancer in each of the 50 US states in relation to the average percentage of women in each state undergoing annual PAP smear screening.

   b. A study that compares the prevalence of back pain among current members of the plumbers and pipe-fitters union with that of current members of the bakers and confectionary union.

   c. A study that evaluates the relationship between breast cancer and a woman's history of breastfeeding. The investigator selects women with breast cancer and an age-matched sample of women who live in the same neighborhoods as the women with breast cancer. Study subjects are interviewed to determine if they breastfed any of their children.

   d. A study of the relationship between exposure to chest irradiation and subsequent risk of breast cancer that was begun in 2005. In this study, women who received radiation therapy for postpartum mastitis (an inflammation of the breast that occurs after giving birth) in the 1940s were compared to women who received a non-radiation therapy for postpartum mastitis in the 1940s. The women were followed for 50-60 years to determine the incidence rates of breast cancer in each group.