What will you learn from this course?

By the end of this course you should be able to:

- understand attributes of geographic data, georeferencing systems and the importance of spatial autocorrelation in spatial data analysis
- display a map within the GIS package ArcView and perform basic query operations and calculations
- plot point data using a GIS package and create a kernel density surface, based on plotted point data
- describe the structure of a fixed-effects model of disease count data and interpret the regression coefficients from a mixed-effects model
- select, apply and interpret the results of regression methods for the analysis of case–control and cohort studies, using appropriate computer software
- plan a strategy of analysis for an epidemiological data set, using an appropriate choice of methods and statistical models
- describe the effects of correlated data on epidemiological analysis, and the use of statistical methods which take account of such correlations
- review and summarize information from many studies using meta-analysis
- estimate the risk attributable to an exposure in a population.

Although this course has been designed to avoid unnecessary mathematical detail, a proper understanding of the methodology requires the use of some mathematical formulae. No knowledge of calculus will be required.

Course structure

The course consists of 14 units of study, all of which you should complete. They make up the following three modules.

Module 1: Investigation of Spatial Patterns of Animal Disease

In this module (Units 1–4) geographical information systems are introduced, based on the ESRI ArcView software, and computer-based activities will guide you through the process of producing maps from spatial data. You will learn how to undertake exploratory analyses and statistical modelling of spatial data, using DynESDA and the ESRI Spatial Analyst extension tools.

Module 2: Advanced Methods of Statistical Analysis Part I

The second module (Units 5–8) is developed as Computer-Assisted Learning (CAL) material. The first unit of this module demonstrates how to review and summarize the framework and process of generalized linear regression models, using logistic regression as an illustration. Units 6 and 7 are about classical analysis of matched case–control studies. In Unit 8, different sampling schemes for case–control studies and the exposure odds ratio estimates are described.

Module 3: Advanced Methods of Statistical Analysis Part II

The final module (Units 9–14) is also developed as CAL units. This module builds on the advanced statistical methods that were illustrated in Module 2. Unit 9 discusses the Poisson distribution, reviews Poisson regression and over-dispersion within a Poisson model. In Unit 10 you will learn about further issues for Cox regression and in Unit 11 about strategies for obtaining regression models that adequately describe the data. Unit 12 deals with how to conduct analyses of data based on correlated observations and Unit 13 introduces statistical approaches to meta-analysis and systematic review. The last unit focuses on attributable risk and on how to estimate the fraction of disease occurrence attributable to an exposure in a population.

Workbook for Modules 2 and 3

In addition to the ten CAL units that make up Modules 2 and 3, there are seven practical computer sessions using the statistical package R. During these Workbook sessions, from