MODULE SPECIFICATION

Academic Year (student cohort covered by specification)	2023-2024
Module Code	2465
Module Title	Analysis of Hierarchical and Other Dependent Data
Module Organiser(s)	Prof Linda Sharples; Prof James Carpenter
Faculty	Epidemiology & Population Health
FHEQ Level	Level 7
Credit Value	**CATS:** 15 **ECTS:** 7.5
HECoS Code	101031 : 101034
Term of Delivery	Term 2
Mode of Delivery	For 2023-24 this module will be delivered by predominantly face-to-face teaching modes. Where specific teaching methods (lectures, seminars, discussion groups) are noted in this module specification these will be delivered by predominantly face-to-face sessions. There will be a combination of live and interactive activities (synchronous learning) as well as self-directed study (asynchronous learning)
Mode of Study	Full-time
Language of Study	English
Pre-Requisites	Students must have masters level modules in linear regression and generalised linear models (or equivalent), and be familiar with parallel group randomised trial design and the principles of confounding. For students on the intensive MSc Medical Statistics programme, pre-requisites are Regression (term 1), Clinical Trials (term 1) and 2462: Statistical Models for Discrete Outcomes (term 2). and be familiar with issues in clinical trial design.
Accreditation by Professional Statutory and Regulatory Body	None

Module Cap (Maximum number of students)	35 (numbers may be capped due to limitations in facilities or staffing)
Target Audience	This module is intended for students who wish to acquire applied statistical skills and understanding of the associated theoretical concepts for the analysis of dependent, specifically hierarchical and longitudinal, continuous and categorical data.
Module Description	The module will cover statistical methodology for modelling dependent data using both likelihood and moment-based methods. Practicals will show how the methods can be applied to analyse and interpret dependent data. In particular, the module focus is on modelling measurements that dependent because they are clustered or because they are repeated measurements over time on the same subject. In this module, we begin by considering normally distributed data, and two level hierarchies, before moving to consider categorical data. Both likelihood and moment-based (marginal) methods will be covered. Methods for missing data will be briefly mentioned as necessary but are not the principal focus of any sessions.
Duration	5 weeks at 2.5 (scheduled sessions only on Mon, Tue; Wed self-directed time
Timetabling slot	Slot D1
Last Revised (e.g. year changes approved)	June 2023

Programme(s) This module is linked to the following programme(s)	**Status** *(Compulsory/Recommended Option)*
MSc Medical Statistics	Recommended

Module Aim and Intended Learning Outcomes

Overall aim of the module
The overall module aim is to: • give students a practical understanding of how to analyse hierarchical, longitudinal and other dependent continuous and discrete data, motivated by examples arising in clinical trials and observational studies.

Module Intended Learning Outcomes
Upon successful completion of the module a student will be able to: 1. Recognise dependent data and have an understanding of why models for independent data are not appropriate for their analysis 2. Demonstrate an understanding of the theoretical basis of the most commonly adopted methods for the analysis of hierarchical, longitudinal and other dependent continuous and categorical data 3. Make appropriate practical use of selected techniques for the analysis of dependent data with appropriate statistical software (Stata)

Indicative Syllabus

Session Content
The module will cover the following topics: • The concept and consequences of dependence • Fixed and random effects in analysis of variance and variance components models for continuous data • Longitudinal data analysis including likelihood based hierarchical models for continuous outcome data • Marginal likelihood based model models for hierarchical and longitudinal continuous data structures • Assessment, choice and impact of covariance structure • Introduction to generalised estimating equations • Extension of likelihood based and generalised estimating equation based methodology to the analysis of dependent categorical data • Interpretation of conditional and marginal models • Checking modelling assumptions • Dependent data modelling using Stata

Teaching and Learning

Notional Learning Hours		
Type of Learning Time	**Number of Hours**	**Expressed as Percentage (%)**
Contact time	48	32
Directed self-study	28	19
Self-directed learning	30	20
Assessment, review and revision	44	29
Total	**150**	**100**

Student contact time refers to the staff facilitated face-to-face time allocated to teaching through lectures and computer practicals.

The division of notional learning hours listed above is indicative and is designed to inform students as to the relative split between interactive (online or on-campus) and self-directed study.

Teaching and Learning Strategy

Guidance Note: Please describe the teaching and learning strategy for the module including the formative tasks for the module.

The module is presented in sessions.

Sessions consists of a 1h lecture followed by a 1.5h computer practical.

Immediately following lectures, practical sessions ensure that students have the opportunity to apply the concepts and methods covered by lecture content.

The practical sessions provide students with “hands on” experience in analysing and interpreting data, using data sets drawn from research work of staff in the faculty or from other sources.

Students are provided with detailed solutions to the tasks set in practical sessions, enabling them to check their understanding of the material.

The module may conclude with a guest lecture from a statistician applying the techniques to tackling current public health challenges.

The assessment task, which comes towards the end of the module, is the point at which students demonstrate a consolidation of their learning across the whole module.

Teaching and Learning Strategy
Practicals will be in STATA: all students will need access to this package.

Assessment

Assessment Strategy
The assessment for this module has been designed to measure student learning against the module intended learning outcomes (ILOs) as listed above. The assessment will consist of analysis of a dataset which we will provide, and production of a written report. The grade for summative assessment(s) only will go towards the overall award GPA.

Summative assessment			
Assessment Type	**Assessment Length (i.e. Word Count, Length of presentation in minutes)**	**Weighting** (%)	**Intended Module Learning Outcomes Tested**
Coursework	4-page report	100	1 – 3

Resitting assessment
Resits will accord with the LSHTM's Resits Policy Resit/deferred/new attempts - The task will be to carry out the analysis of a new dataset. The next assessment deadline will be during mid/late September of the current academic year.

Resources

Indicative reading list *(if applicable)*
Guidance note: Please list up to 12 core texts and sources for the module.
References on which the course is based

Background reading:
Introducing Multilevel Modelling, by Ita G. Kreft & Jan de Leeuw. Publisher: SAGE Publications, Ltd

Broader reading:
1. Rabe-Hesketh, S. and Skrondal, A. (2012) Multilevel and Longitudinal Modeling Using Stata, 3rd Edition. Stata Press.
2. Snijders, T. and Bosker, R. (1999) Multilevel Analysis SAGE Publications Ltd.
3. Verbeke, G. and Molenberghs, G. (2000) Linear Mixed Models for Longitudinal Data. Springer Verlag.
4. Fitzmaurice, G.M., Laird, N.M., and Ware, J.H. (2011) Applied Longitudinal Analysis. 2nd edition. John Wiley and Sons, New York.

Other important references
1. Diggle, P.J.,Heagerty, P., Liang, K.-Y. and Zeger, S.L. (2002) Analysis of Longitudinal Data, Second Edition Oxford University Press.
2. Dwyer, J.H., Feinleib, M., Lippert, P. and Ho_meister, H. eds (1990) Statistical Methods for Longitudinal Studies of Health. Oxford University Press.
3. Fitzmaurice, G.M., Laird, N.M. and Ware J.H. (2004) Applied Longitudinal Analysis. Wiley.
4. Goldstein, H. (2011) Multilevel Statistical Models, Fourth Edition. Arnold, London.

<u>Other resources</u>
Guidance note: Please list the other study resources for the module.
STATA will be made available to all students. A set of reading notes with worked examples will be available on Moodle.

Teaching for Disabilities and Learning Differences

The module-specific site on Moodle provides students with access to lecture notes and copies of the slides used during the lecture prior to the lecture (in pdf format). All lectures are recorded and made available on Moodle as quickly as possible. All materials posted up on Moodle areas, including computer-based sessions, have been made accessible where possible.

The LSHTM Moodle has been made accessible to the widest possible audience, using a VLE that allows for up to 300% zoom, permits navigation via keyboard and use of speech recognition software, and that allows listening through a screen reader. All students have access to "SensusAccess" software which allows conversion of files into alternative formats.

For students who require learning or assessment adjustments and support this can be arranged through the Student Support Services – details and how to request support can be found on the LSHTM Disability Support pages.