

Medical Statistics 1 (STAT0014)

In-Course Assessment I (2023-24)

Basic Information

- This is a group assessment: this means that all members of the group are responsible for contributing to the overall result and you will all be given the same mark. The final group allocation can be found on the Moodle page of the course.
- The release date for this assessment is 17:00 (UK time) on Tuesday, 28th November 2023.
- Your ICA report should be uploaded onto Moodle by **THURSDAY 7th DECEMBER 2023 by 5pm**. For the work to be submitted successfully **every group member must press the submit button.**
- There are three questions in this ICA. You will need to answer all questions and produce an ICA report. Each group needs to submit the ICA report **as a single PDF file**.
- Name your report as: 'group_id_report.pdf'. For example, if you are in group 'A' the file should be named **group_A_report.pdf**.
- The full report must be **no more than 9 A4 sides** (including tables and figures but excluding the appendix and the ownership declaration) based on single spaced text with font size of 11.
- Tables for all questions must be in a form suitable for presentation of results in a report, which **will not be the same** (or require as much detail) as that given by Stata. Results need to be formatted carefully selecting the appropriate information to present. (Also note that $P=0.000$ should be changed to (e.g.) $P < 0.001$).
- **Do not include your names on your work.** Please include only your **student numbers** (SRN, as printed on the front of your UCL identity card).
- This ICA contributes **15% of your overall mark for this module**.
- You will receive a *provisional* grade for this ICA towards the end of January 2024 (*grades are provisional until confirmed at the Statistics Examiners' Meeting in June 2024*).
- In preparation for this assessment, please ensure that you are familiar with the Department of Statistical Science's [guidance on academic integrity](#). When submitting your work, you will be required to make a declaration that you have read and understood this guidance.
- Parts of your submission may be scanned using similarity detection software. If any breach of the assessment regulations is suspected, it will be investigated in accordance with UCL's [Student Academic Misconduct Procedure](#).
- To facilitate anonymous marking, you should not write your name anywhere on your work, including in file names or file descriptions requested as part of the submission process.
- You must only submit your work via the designated portal in Moodle. If you try to submit via email or any other channel this will not count as a submission and will not be marked.

- AI tools cannot be used.
- There are strict, non-negotiable [penalties for late submission](#), which for coursework are as follows.
 - Up to 2 working days late: deduction of 10 percentage points, but no lower than the pass mark.
 - 2-5 working days late: capped at the pass mark.
 - More than 5 working days late: mark of 1.00%.
- If the module lead becomes aware of a significant technical issue or outage affecting Moodle during the assessment, a message will be circulated to explain what has happened and the steps being taken to mitigate the issue. If you do not receive notification of a more widespread issue and you experience technical difficulties, you should refer to the [Help & Support](#) resources provided by UCL's central IT service. However, last-minute technical issues will not be considered as valid grounds for missing the deadline, so ensure that you leave plenty of time to prepare, upload and check your submission.
- Non-submission (in the absence of valid extenuating circumstances) will mean that your mark for this component is recorded as 0.00% and you will be deemed to have made an attempt.
- Extensions to the submission deadline can only be granted where a student has been issued with a SoRA or has made a valid claim for extenuating circumstances. The standard extension length for this assessment type is one week.
 - If you have a [SoRA](#), individual extensions to the submission deadline will not be applied automatically due to the nature of the assessment (group work). If a group member with a SoRA chooses to activate their extension, then the extension will be applied to the entire group and be reflected in the deadline displayed in the submission portal. If you have chosen to activate your extension and you think that your SoRA adjustment has not been applied, please contact the module lead at the earliest opportunity.
 - [Extenuating circumstances](#) are handled by your parent department and all claims should be submitted via [Portico](#). Depending on the nature and severity of the circumstances, an alternative type of mitigation to a deadline extension may be considered more suitable.

Working in a group

- You must work within your group. You are expected to meet to discuss how to approach the work as a team at the beginning of the ICA period and later on to discuss and prepare the submission. For example, you may decide that you all attempt the work, then discuss and each produce different parts of the final report.
- Each group needs to submit a single, collective piece of work on behalf of the whole group. One of the objectives of the ICA is to prepare you to work in a team, where each member takes responsibility for the success of the work. You may decide to name a "team captain" and they can take charge of the overall project or divide up the work uniformly among the members.
- If there are problems with team-members not engaging, please contact the lecturer and report the incident.

Declaration of ownership

In addition to the answers to the ICA questions, all groups **must include one additional page (the last page in the report)** in which:

- 1) You briefly describe the **individual contribution of each member of the group**.
 - Each one of you will need to sign this declaration, signifying that you are happy with the representation of your participation in the project.
 - Note that this page will not be used to specify different marks for each member of the group. If one member does not engage with the group, you are encouraged to report this. In very exceptional circumstances, the marks may be adjusted accordingly.
- 2) You include a **Declaration of ownership**, i.e., the following text:

"We are aware of the UCL Statistical Science Department's regulations on plagiarism for assessed coursework. We have read the guidelines in the student handbook and understand what constitutes plagiarism. By submitting the ICA report as detailed in the text of the ICA, we hereby affirm that the work has entirely been carried out by us"

A copy of the guidelines on plagiarism and collusion are available in the student handbook and on the ICA section of the course. Please ensure you have read these. In submitting your report online, you will also be prompted to affirm that you have read the guidelines on what constitutes plagiarism and collusion and that the work being submitted is entirely your own.

Question 1 (55 marks)**Background**

Between 6% and 10% of visits to primary-care practitioners (GPs) in the UK result in a referral to hospital for a face-to-face meeting with a specialist doctor. The process of referral requires effective communication between all parties involved (GP, specialist and patient), and deficiencies can lead to a range of problems. Use of video-conferencing technologies to allow a virtual meeting between a patient and GP (both at the general practice) and the specialist at the hospital, has been suggested as a possible alternative to the usual system of a face-to-face meeting. It is hoped that this alternative approach would be more efficient and therefore reduce the need for patients to have a subsequent face to face follow up appointment with the specialist at the hospital.

The usefulness of these video consultations was considered in a randomised trial comparing against usual hospital appointments. The trial was a **multicentre randomised parallel group design with one urban (London) and one rural site (Shrewsbury)**. The primary outcome was whether or not the patient required a subsequent follow up hospital appointment within 6 months of their trial consultation (a binary outcome).

Data

The Stata version of the data for this trial is the **dataset_q1_x.dta** file, where x corresponds to the **individual dataset matched to your group ID**. The data can be found on the course Moodle page in the ICA section (folder 'Data') – please ensure you download the file relevant to your group ID.

The variables in the dataset and their codings are:

Randomised treatment group:

randgrp Randomised group (1=video_consult, 0=standard)

Measured at baseline:

site 0=London 1=Shrewsbury
 specgrp specialty: 1= Orthopaedics, 2=Urology, 3=ENT, 4=Gastroenterology, 5= other
 age age at randomisation (in years)
 gender 0=male, 1=female
 ethngrp ethnic group: 1=white, 2= Black Caribbean/ African, 3=Indian/ Pakistani/ Bangladeshi, 4= Chinese, 5=Other Asian, 6=Other
 dusoi Baseline Duke Severity of Illness score (0=low severity, 100=high severity)
 SF12phy0 Baseline SF12 physical health score (0 to 100=optimal physical health) [measured for adults only]

Measured at follow up:

appointment follow up appointment required? (**primary outcome**)
 (0=no appointment, 1=appointment)

Protocol deviations information:

withdrawn patient withdrew (dropped out) from trial during follow up (1=yes, 0=no)
 compliance compliance with trial intervention (1=compliant, 0=noncompliant)

Note that although 3170 patients were considered for the trial, 46 were found to be ineligible and 1135 failed to provide consent. The dataset contains data only for the 1989 patients who were eligible for the trial **and** were randomised.

Trial analysis plan

The brief statistical analysis plan for the trial requires:

- 1) A **consort diagram**. [4 marks]
- 2) A **table** of baseline characteristics. [5 marks]
- 3) An unadjusted (complete case, ITT) analysis for the primary outcome (reporting both absolute and relative effect measures). [8 marks]
- 4) An analysis of the primary outcome adjusting for **site, gender, age and specialty** (reporting a relative effect measure). [4 marks]
- 5) An investigation of whether the treatment effect for the primary outcome **differs by study site** (reporting results using a relative effect **measure**). Do not adjust for gender, age and specialty in this analysis. [11 marks]

Your task

- A) Carry out analyses to **address each requirement of the plan above**. [total 32 marks, maximum marks for individual sections are shown above]
 - You should present the results using clearly titled and labelled tables or figures. For 2) - 5) you should include tables of relevant descriptive statistics.
 - After each numbered analysis (1-5 of the plan) you should include an interpretation (**maximum 4 sentences presented using bullet points**).
- B) Include an annotated Stata do-file in an Appendix at the end of your report showing all commands used to obtain the results reported (STATA output should not be included). [3 marks]
- C) Based on the analysis in part A, make a brief overall interpretation about the usefulness of the video-consultation intervention, commenting on whether this is as expected. [4 marks]
- D) Identify two problems with this study and describe any analyses (sensitivity or other) which might be used to investigate or address each problem (maximum 250 words in total for both problems). [10 marks]
- E) Suppose that in addition to knowing **whether** a patient had a follow-up appointment within 6 months since the initial consultation, you also have information on the **exact date** of the initial consultation and the date of the follow-up consultation (if it took place). In the light of this additional information describe how you would perform an alternative analysis to the unadjusted analysis in 3) above. State any assumptions of your alternative analysis (maximum 150 words). [6 marks]

Question 2 (15 marks)

Background

Helicopter emergency medical services are sometimes used for rapid interfacility transportation of traumatically injured patients, but the effect on clinical outcomes for patients is debatable. As helicopter transportation is an expensive and often limited resource, a study was conducted to investigate its effectiveness in improving patient outcomes compared with ground ambulance emergency medical services.

Data

Data were collected from 1425 patients who experienced traumatic injuries in the Alps region over the period 2000-2015 and were transferred to nearby hospitals. The outcome of interest was in-hospital death (binary outcome) following transportation. Data were available on mode of transportation (Ground Ambulance/ Helicopter), severity of the traumatic injury (categorised as low/high) and patient age (in years).

The Stata version of the data for this study is the **dataset_q2_x.dta** file which can be found on Moodle page in the ICA section (folder 'Data'), where x corresponds to **the dataset matched to your group ID.**

The variables in the dataset and their codings are:

died	in-hospital death; 0=no, 1=yes
mode	mode of transportation; 0=ground ambulance, 1=helicopter
severity	severity of injury; 0=low, 1=high
age	age at the time of injury (in years)

Your task

- Perform an exploratory investigation to examine whether age and severity of injury may be confounders in the association between in-hospital death and mode of transportation. This should include tables and/or figures with appropriate interpretation and conclusion. [8 marks]
- Write down a suitable model to establish whether helicopter transportation is associated with improved in-hospital survival or not (compared with ground ambulance). Justify this model based on your exploratory investigation, fit the model and interpret the results. [5 marks]
- Include an annotated Stata do-file in the Appendix (same as in Q1) showing all commands used to obtain the results reported (STATA output should not be included). [2 marks]

Question 3 (20 marks)**Background**

Chronic headaches affect the quality of life for its sufferers and are an important public health challenge due to their associated costs. Available medications may help some people, while others look for alternative approaches. A parallel group randomised trial was conducted to determine the effect of **acupuncture therapy vs. no acupuncture** therapy on headache in patients with chronic headaches. The primary outcome was a headache severity score. Patients were assessed at baseline and at 12 months post-randomisation.

Data

The Stata version of the data for this study is the **dataset_q3_x.dta** file which can be found on Moodle page in the ICA section (folder 'Data'), where x corresponds to **the dataset matched to your group ID.**

The variables in the dataset and their codings are:

Randomised treatment group:

group randomised group (1=acupuncture; 0=no acupuncture)

Measured at randomisation:

age age at randomisation (in years)

sex 0=male, 1=female

hs0 headache severity score at randomisation (0=low severity, 100=high severity)

Measured at follow up:

hs12 headache severity score at 12 months post randomisation (0=low severity, 100=high severity)

The study statistician was on long-term leave and one of the investigators, eager to see some results, performed a quick analysis by comparing the **average headache severity scores at 12 months between the two groups.** A suspicious colleague re-analysed the data, suggesting that it is better to compare the **average of differences** between headache severity at the time of randomisation and at 12 months, between the two groups. From these results the investigators were puzzled as to whether or not there really was evidence that acupuncture works. They consult you for help.

Your task

- A) Perform the two analyses mentioned above, present and interpret each result (including appropriate tables/figures). [8 marks]
- B)** Based on the analyses in A) and/or additional analyses you think appropriate, state your conclusion regarding the **effect of acupuncture**. Explain to the investigators any differences between the treatment effect estimates obtained from different methods and justify your preferred approach for this analysis (you don't need to consider whether age and sex are treatment effect modifiers). [10 marks]
- C) Include the annotated Stata commands in the Appendix (same as in Q1 and Q2) showing all commands used to obtain the results reported (STATA output should not be included). [2 marks]

The full report must be **no more than 9 A4 sides** (including tables and figures but excluding the Appendix and the Ownership Declaration) based on single spaced text with font size 11.