MSc Degree Examinations 2019/20

Department:
Health Sciences

Title of Exam:
Measurement in Health and Disease

Time Allowed:
24 hours
(30 hours for students entitled to extra time)

Allocation of Marks:
Marks are indicated on the exam paper

Instructions for Candidates:

Set up a word document on your computer. Please note you do *not* need to include a coversheet. You should *not* include any reference to your identity. You should include only your exam candidate number in the header. Please **answer all the below questions for parts A, B and C**. There is no need to repeat the questions in your document; just use the question number (and letter for subsidiary questions where appropriate, i.e. 1a), 1b)…) for reference.

If a question is unclear, answer the question as best you can, and note the assumptions you've made to allow you to proceed.

Submission Information:

Please submit your anonymous exam as a pdf via the VLE submission point by the deadline on **Tuesday 16 June 2020 by 09:30.** The submission deadline for students with extra time is **Tuesday 16 June 2020 by 15:30**.
To submit your work to go the Health Research Methods VLE page. Click on **Assessment** on the side menu, click on **Assessment Submission.** On this page you will find the **'Submit your exam here'** tab which you should click on and follow the instructions regarding submission.

Should you experience problems with the VLE, there is an option to email your exam paper to the Student Assessment office at dohs-assessments@york.ac.uk. Note this option should only be used if submission via the VLE is not possible for you for technical reasons.

EXAMS SUBMITTED PAST THE SUBMISSION DEADLINE WILL NOT BE MARKED AND WILL BE CONSIDERED A NON-SUBMISSION

A note on Academic Integrity

We are treating this online examination as a time-limited open assessment, and you are therefore permitted to refer to written and online materials to aid you in your answers.

However, you must ensure that the work you submit is entirely your own, and for the whole time the assessment is live you must not:

- communicate with departmental staff on the topic of the assessment;
- communicate with other students on the topic of this assessment;
- seek assistance with the assignment from the academic and/or disability support services, such as the Writing and Language Skills Centre, Maths Skills Centre and/or Disability Services (the only exception to this will be for those students who have been recommended an exam support worker in a Student Support Plan. If this applies to you, you are advised to contact Disability Services as soon as possible to discuss the necessary arrangements);
- seek advice or contribution from any third party, including proofreaders, friends, or family members.

We expect, and trust, that all our students will seek to maintain the integrity of the assessment, and of their award, through ensuring that these instructions are strictly followed. Failure to adhere to these requirements will be considered a breach of the Academic Misconduct regulations, where the offences of plagiarism, breach/cheating, collusion and commissioning are relevant - see AM.1.2.1" *(Note this supersedes section 7.3 of the Guide to Assessment).*

PART A   (30 marks)

1. **A retrospective diagnostic accuracy study investigated predictors of pre-term birth and stillbirth from maternal pre-diabetic indicators during pregnancy, such as the oral glucose tolerance test (OGTT) and antenatal random glycaemia (ARG). Results included the following receiver operating characteristic (ROC) curves.**
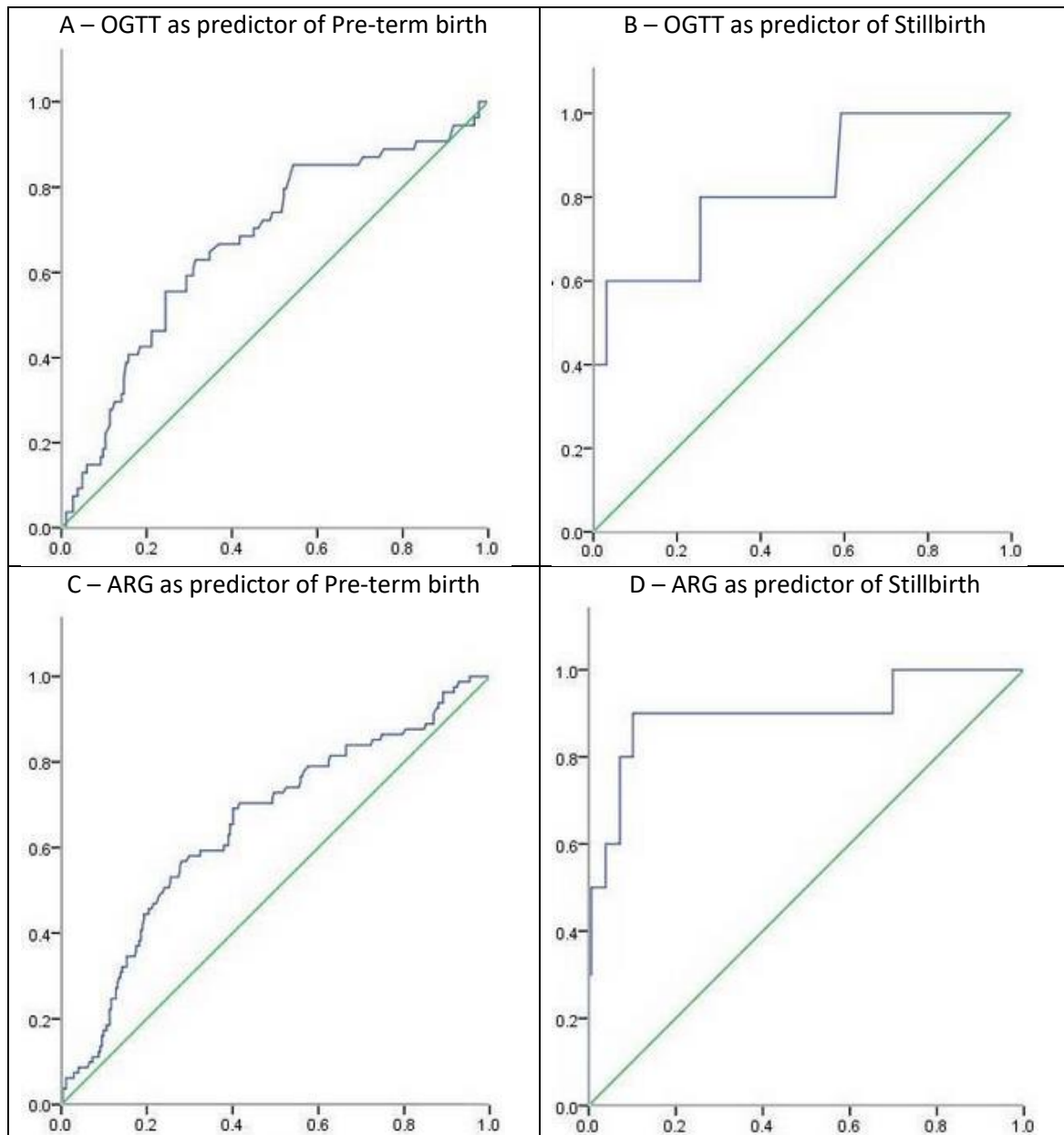
A – OGTT as predictor of Pre-term birth

B – OGTT as predictor of Stillbirth

C – ARG as predictor of Pre-term birth

D – ARG as predictor of Stillbirth

Select the <u>single best answer</u> from the response options for each question, note the corresponding letter in your answer document <u>AND</u> justify your answer briefly. There are no penalties for selecting an incorrect answer.

1a. What can be said for the ability of ARG and OGTT to predict birth outcomes?   (3 marks)

A   ARG is a better predictor of pre-term birth, but similar to OGTT for stillbirth

B   ARG is a worse predictor of pre-term birth, but similar to OGTT for stillbirth

C   ARG is a better predictor of stillbirth, but similar to OGTT for pre-term birth

D   ARG is a worse predictor of stillbirth, but similar to OGTT for pre-term birth

E   It is impossible to tell from the graphs alone

1b. What do ROC curve values to the right of the diagonal line indicate (e.g. top right of panel A)?   (4 marks)

A   Diagnostic accuracy is not statistically significant at that point

B   The diagnostic test predicts pre-term births better than chance at that point

C   The diagnostic test predicts non-pre-term births better than chance at that point

D   The diagnostic test is worse than average at that point

E   The estimate of diagnostic accuracy is unreliable at that point

1c. What could be the reason for the 'steps' of the plot in panels B and D to be bigger than in panels A and C?   (3 marks)

A   Cut-off values for OTT and ARG are more likely to be clustered for stillbirths than pre-term births

B   It is likely that there were fewer stillbirths than pre-term births in the study sample

C   OGTT and ARG are better predictors of stillbirths than pre-term births

D   OGTT and ARG are worse predictors of stillbirths than pre-term births

E   Sensitivity and specificity can be expected to be more stable for stillbirths than pre-term births

2. In an emergency care setting, the comparison between measured (co-oximetry) and calculated (calc) central venous oxygen saturation (ScvO$_2$) yielded the following 'limits of agreement' plots for all patients in the study and those with septic or non-septic shock.
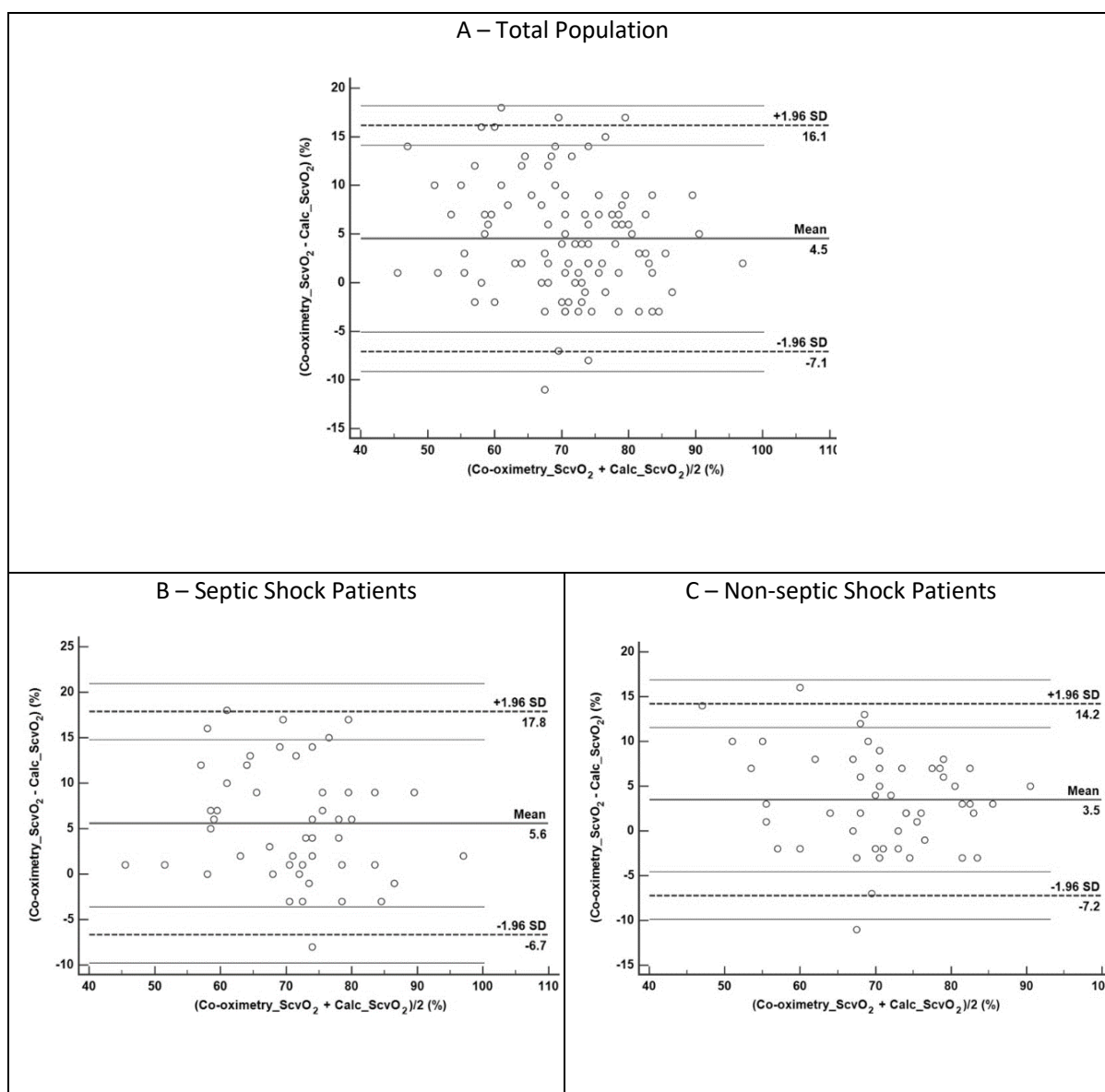
A – Total Population

B – Septic Shock Patients

C – Non-septic Shock Patients

Select the <u>single best answer</u> from the response options for each question, note the corresponding letter in your answer document <u>AND</u> justify your answer briefly. There are no penalties for selecting an incorrect answer.

2a. Is there any evidence for systematic bias between the two methods of assessing ScvO$_2$?

(4 marks)

A   It is impossible to tell from the graphs

B   There are no notable differences in systematic bias

C   Yes, calculated ScvO$_2$ is generally larger than co-oximetry ScvO$_2$

D   Yes, co-oximetry ScvO$_2$ is generally larger than calculated ScvO$_2$

E   Yes, there is some evidence for bias, but this is not consistent between septic and non-septic shock patients

2b. How would you interpret the differences in agreement between septic and non-septic shock patients?
(3 marks)

A   Agreement is worse for non-septic patients, as more data points lie outside the limits of agreement

B   Agreement is worse for non-septic patients, as the lower limit of agreement is lower

C   Agreement is worse for septic patients, as the limits of agreement are wider

D   Agreement is worse for septic patients, as there is additional bias over the range of possible ScvO$_2$ values

E   There is no notable difference in agreement between the two groups

2c. In the total population, is there bias in the agreement over the range of ScvO$_2$ values?
(3 marks)

A   No, agreement does not appear to be related do the magnitude of ScvO$_2$

B   Yes, agreement is worse for higher ScvO$_2$ values

C   Yes, agreement is worse for lower ScvO$_2$ values

D   Yes, agreement is worse for mid-range ScvO$_2$ values

E   Yes, agreement is worse for outlying ScvO$_2$ values at either end

3. Decide which statistic is appropriate in the following scenarios.

Select the single best answer from the response options for each question, note the corresponding letter in your answer document **AND** justify your answer briefly. There are no penalties for selecting an incorrect answer.

3a. Which statistic would you use to describe the agreement of measuring the presence of lateral downsloping (present/absent) on the shoulder MRI of 140 patients with rotator cuff tears between five radiologists? **(4 marks)**

A   Area under the curve

B   Cohen's kappa

C   Cronbach's alpha

D   Eigenvalue

E   Intra-class correlation coefficient

3b. Which statistic would you use to describe the agreement of measuring the lateral acromial angle (in degrees) on the shoulder MRI of 140 patients with rotator cuff tears between five radiologists? **(3 marks)**

A   Area under the curve

B   Cohen's kappa

C   Cronbach's alpha

D   Eigenvalue

E   Intra-class correlation coefficient

3c. Which statistic would you use to describe the reliability of measuring the subacromial space (in millimetres) on the shoulder MRI of 140 patients with rotator cuff tears, assessed twice by the same radiologist? **(3 marks)**

A   Area under the curve

B   Cohen's kappa

C   Cronbach's alpha

D   Eigenvalue

E   Intra-class correlation coefficient

PART B   (30 marks)

Questions relating to paper: Manual muscle testing and hand-held dynamometry in people with inflammatory myopathy: An intra- and interrater reliability and validity study. (Baschung Pfister et al. 2017)

4. Table 3 presents the intra-class correlation coefficients (ICC), standard errors of measurement (SEM) and smallest detectable change (SDC), the latter expressed as statistics of raw measurements as well as of percentage change. How are these statistics defined in general and what do they represent in the context of the data presented for this study? Interpret and compare the intra- and interrater reliabilities for elbow flexion and ankle extension.

    (10 marks)

5. Table 5 presents a number of Cohen's kappa values. How is this statistic defined in general, and what do the values represent in the context of this study? Why do you think a weighted version of kappa was used, and do you agree the use of kappa is appropriate here? Interpret and compare the intra- and interrater reliabilities for elbow flexion and ankle extension.

    (10 marks)

6. In Figure 3 and Figure 4, agreement between measurements is represented by limits of agreement plots (or Bland-Altman plots). In general, what is a Bland-Altman plot, and what is the purpose of drawing such plots? How have they been used in this study? Interpret and compare the intra- and interrater reliabilities for elbow flexion and ankle extension.

    (10 marks)

Part C   (40 marks)

Questions relating to paper: Criterion Validity of the Yale-Brown Obsessive-Compulsive Scale Second Edition for Diagnosis of Obsessive-Compulsive Disorder in Adults. (Castro-Rodrigues et al 2018)

7. In the paper, the authors refer to convergent and divergent validity. Define what these terms mean in general and in the context of this study. How is each type of validity assessed in the paper, and what would desirable outcomes (high levels of validity) look like? To what extent have these validities been demonstrated in your opinion?   **(10 marks)**

8. The authors conducted a factor analysis and reported results in Figure 1. Define what a factor analysis refers to in general and why it was conducted in this study. How did the authors decide on the number of factors, and would you agree with the number selected? In the table presented in Figure 1, interpret factor loadings for items 1, 3 and 4 in the context of the items and factors.   **(10 marks)**

9. In the text on page 5 and in Table 2, the authors report Cronbach's alpha for the PY-BOCS-II. What does this statistic represent and how can it be interpreted in general? How was it used in the context of this study? Interpret the values presented by the authors, contrasting total and sub-scale values, as well as individual item-deletion alpha values for the total and OCD-subset sample.   **(10 marks)**

10. In Table 4 the authors present sensitivity and specificity. What do the terms sensitivity and specificity refer to in general and in the context of this table? How was the cut-off of 13 points decided? In Figure 2, there are two additional ROC curves for matched controls and blinded assessment. Why do you think these additional analyses were conducted, and how would you *expect* results to differ from the main analysis? **(10 marks)**

End of paper