



AMREF INTERNATIONAL UNIVERSITY
SCHOOL OF PUBLIC HEALTH
DEPARTMENT OF COMMUNITY HEALTH
BSC HEALTH SYSTEMS MANAGEMENT AND DEVELOPMENT /
HEALTH RECORDS AND INFORMATION MANAGEMENT
END OF SEMESTER FINAL EXAMINATION AUGUST 2026

UNIT CODE: HRI 209
HEALTH CARE DATA

UNIT NAME: ANALYSIS AND INTERPRETATION OF

DATE: 7th August, 2026

TIME: Two Hours

Start: 5:30pm

Finish: 7:30 pm

INSTRUCTIONS

1. This exam is marked out of 70 marks
2. This Examination comprises TWO Sections
Section A: Compulsory Question (30 marks)
Section B: Long Answer Questions (40 marks)

INSTRUCTIONS: Answer Question ONE and any other TWO Questions.

Section A: Answer all questions (30Marks)

Question one (30 Mks)

1. A patient arrives at a health facility with a temperature of 39°C, blood pressure of 150/100 mmHg, and a positive malaria test.
 - a) Differentiate between data and information using the above example. (4 Marks)
 - b) Explain three ways a hospital data analyst can support clinical decision-making using health data. (3 Marks)
2. Explain three importance of ICD coding in standardizing healthcare data. (3 Marks)
3. Identify three characteristics that distinguish between primary and secondary sources of health care data. (3 Marks)
4. Quality improvement analysts at City General Hospital are evaluating the effectiveness of the Outpatient Triage system in predicting subsequent healthcare utilization.

Specifically, the hospital board wants to know if the initial triage acuity level of a patient is associated with the likelihood of being readmitted within 30 days of discharge. A chi-square test was conducted.

Chi-Square Tests

Pearson Chi-Square = 12.56

df = 2

p-value = 0.002

- a) State the hypothesis for the above study. (4 Marks)
- b) Can the hospital quality team conclude that being assigned a high-acuity triage level directly causes a higher readmission rate? (6 Marks)
5. State four ethical principles that should guide the sharing of healthcare data. (4 Marks)
6. Explain three ways how coding errors can affect health care statistics. (3 Marks)

Section B: Answer any two questions (40Marks)

Question Two (20 Mks)

Read the following clinical data manager's reflection carefully and answer the questions below

"We were under immense pressure to release the oncology clinical dataset to the external university research team by Friday afternoon. The Principal Investigator kept stressing the urgency, arguing that the grant funding depended on this data sharing. When I opened the raw extract, I realized the automated syntax script had failed to drop the medical record numbers (MRNs) and full birth dates because of a formatting glitch. Instead of halting the transfer to rewrite the syntax, I manually deleted a few obvious names in the Excel sheet and sent it via unencrypted email, assuming the university's firewall would protect it. I didn't get explicit consent from the patients for this secondary research use, but the PI assured me that because it's retrospective oncology data, it falls under 'general clinical operations' and we didn't need to bother the IRB or the Data Protection Officer. I felt uneasy about it, especially since some of these patients have rare cancer diagnoses that make them easily re-identifiable even without their names, but I just followed orders to keep the project moving."

- i) Identify and support three distinct themes (6mks)
- ii) Discuss how releasing potentially re-identifiable data without patient consent compromises the pillar of Justice (4mks)
- iii) Outline the compliance steps that were entirely bypassed in this scenario?? (4mks)
- iv) Outline a corrective action plan to mitigate the privacy breach and ensure operational justice for the patients involved (6mks)

Question Three (20 Mks)

1. A Researchers at Kenyatta National Hospital conducted a study on outpatients managing chronic conditions to understand the relationships between lifestyle factors, mental well-being, and treatment compliance. A Pearson correlation analysis was conducted to evaluate these continuous health metrics. The variables analyzed are: Physical Activity Hours (PAH): Weekly hours spent engaging in moderate to vigorous physical activity. Perceived Stress Score (PSS): Standardized score measuring chronic stress (higher scores indicate higher stress). Medication Adherence Score (MAS): Score assessing compliance with prescribed drug regimens. The PSPP correlation output is shown below:

		Physical Activity Hours	Perceived Stress Score	Medication Adherence Score
Physical Activity Hours	Pearson Correlation	1	0.171	0.245
	Sig. (2-tailed)		< 0.001	< 0.001
	N	600	600	600
Perceived Stress Score	Pearson Correlation	0.171	1	0.289
	Sig. (2-tailed)	< 0.001		< 0.001
	N	600	600	600
Medication Adherence Score	Pearson Correlation	0.245	0.289	1
	Sig. (2-tailed)	< 0.001	< 0.001	
	N	600	600	600

- i. State the sample size? (2 marks)
- ii. Identify the variables that have the strongest relationship? justify (3marks)
- iii. Interpret the clinical insight implied by the result in (ii). (3 marks)
- iv. Analyse the relationship between PSS and MAS. (3 marks)
- v. Explain what the relationship in (v) tell you about the variance shared between chronic stress and compliance. (3 marks)
- vi. Based on these statistical findings, propose two targeted clinical interventions or patient support programs that the hospital can implement to improve overall chronic disease management. (6 marks).

Question Four (20 Mks)

Community health researchers at Kenyatta National Hospital conducted a study to assess the risk factors associated with hyperglycemia among adults in an outpatient screening clinic. A binary logistic regression analysis was performed to evaluate the impact of demographic and clinical variables on the likelihood of a patient having hyperglycemia (Yes/No). The dependent variable is Hyperglycemia (1 = Yes, 0 = No). The independent variables are Age

(per 10-year increase), Body Mass Index (BMI > 23 kg/m²), Family history of DM (HDM), and Periodontal status (PD).

The results from the binary logistic regression analysis are shown in the table below:

VARIABLE	B	S.E.	WALD	P-VALUE	OR	95% CI OF OR
AGE (10 YEARS)	0.371	0.071	27.484	< 0.001	1.449	1.261 – 1.665
BODY MASS INDEX (BMI >23 KG/M ²)	0.961	0.204	22.176	< 0.001	2.614	1.752 – 3.900
FAMILY HISTORY OF DM (HDM)	0.558	0.175	10.120	0.001	1.747	1.239 – 2.463
PERIODONTAL STATUS (PD)	0.250	0.121	4.261	0.039	1.284	1.013 – 1.629
CONSTANT	-3.729	0.441	71.646	< 0.001	0.024	-

- i. Write the Logistic regression equation. (4 marks)
- ii. Explain the effect of Age on patient developing hyperglycemia (3 marks)
- iii. Explain the effect of Family History on patient developing hyperglycemia (3 marks)
- iv. Identify the variable that influences the most, Justify (3 marks)
- v. Is the Periodontal status statistically significant? Justify. (3 marks)
- vi. Discuss 2 ways in which these findings can improve the Management of hyperglycemia. (4 marks)