BP703T · B.Pharm Semester VII · Theory Notes

AI in Clinical Applications

Unit III — AI in Personalized Medicine (6 hours)

Unit III (6 hours)

Pills & Python Notes

From-zero teaching standard: every technical term is defined before use.

Syllabus coverage: Unit III (6 hours).

PILLS & PYTHON

Pharmacy • Data • AI • Healthcare

Instagram @pillsandpython · pillsandpython.com · YouTube @PillsAndPython

CHAPTER 7: Precision Medicine — Patient Covariates, Therapeutic Response; Logistic Regression for Disease Risk Prediction

Learning Objectives

After studying this chapter, you will be able to:

1 Define precision medicine, personalized medicine, covariate, therapeutic response, biomarker, risk stratification, and risk score.

2 Explain why the same drug and dose can work differently in different patients.

3 List common patient covariates that change drug response, with examples.

4 Use logistic regression to predict disease risk from patient covariates, with a worked toy example.

5 Interpret a disease-risk probability and place a patient into a risk band.

Definitions

- **Precision medicine:** Treating patients based on their individual characteristics (genes, age, weight, kidney function, lifestyle) instead of giving everyone the same treatment. "The right drug, at the right dose, for the right patient."
- **Personalized medicine:** Mostly the same idea as precision medicine — tailoring treatment to the individual patient. (In exams, you may treat the two terms as synonyms.)
- **Covariate:** Any patient characteristic that can change the outcome and that we measure alongside the main treatment — age, weight, sex, kidney function, other diseases, genes. The word means "a variable that varies along with the patient".
- **Therapeutic response:** How well the treatment works in a patient — did the blood sugar fall? Did the pain go away? Did the infection clear?
- **Biomarker:** A measurable sign in the body that tells something about health or disease — for example, blood sugar level, blood pressure, or a gene variant.
- **Risk stratification:** Sorting patients into risk groups (low, moderate, high) so that high-risk patients get more attention and low-risk patients are not over-treated.
- **Risk score / predicted probability:** The number a model gives for one patient, e.g., "32% chance of developing diabetes in 5 years".
- **Risk factor:** A covariate that increases the chance of disease (e.g., high BMI for diabetes). A protective factor decreases it.

Terminology Table (read this first)

Term	Simple meaning	Pharmacy example
Precision medicine	Treatment matched to the individual	Different warfarin starting doses for different patients
Covariate	Patient characteristic that affects outcome	Age, weight, kidney function, genes
Therapeutic response	Did the treatment work?	Fall in HbA1c after starting metformin
Biomarker	Measurable body signal	Blood glucose, creatinine, a gene test result
Risk stratification	Sorting patients into risk bands	Low / moderate / high diabetes risk
Risk score	Model's probability for one patient	0.40 = 40% predicted risk
Risk factor	Covariate that raises risk	Smoking for heart disease
Protective factor	Covariate that lowers risk	Regular exercise for diabetes
Logistic regression (here)	Same S-curve method as Chapter 5, now predicting disease instead of ADR	Predicting diabetes risk from age and BMI
Reference group	The group everything is compared with	"Non-smokers" as reference for the smoking odds ratio

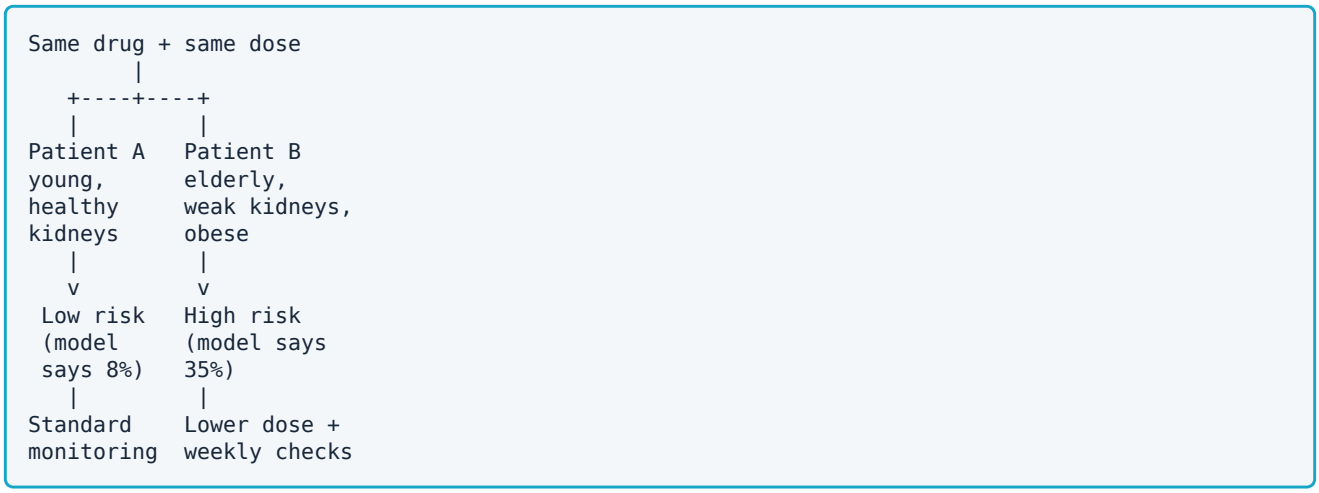
7.1 The core idea: one drug does not fit all

Technical concept: Patients differ in covariates (age, genes, organ function, body size, other diseases). These covariates change both drug response and disease risk. Precision medicine uses measured covariates — often through a prediction model — to individualise decisions.

Simple explanation: The same shirt size does not fit every person. In the same way, the same drug dose does not suit every patient. A young athlete and an elderly kidney patient need different thinking, and a model can put numbers on that difference.

Example: Two patients, both 55 years old, both prescribed the same blood-pressure drug. Patient A is slim with healthy kidneys; Patient B is obese with weak kidneys. Patient B is more likely to get side effects and may need a lower dose or closer monitoring. A risk model turns "Patient B is riskier" into "Patient B has a 35% predicted ADR risk vs 8% for Patient A".

Diagram:



Key takeaway: Covariates explain why responses differ; models convert those differences into numbers a clinician can act on.

7.2 Classification of patient covariates

Category	Examples	How it changes response
Body size	Weight, BMI, body surface area	Bigger body may need bigger dose; obesity changes drug distribution
Age and sex	Elderly, children, male/female	Weaker kidneys/liver in elderly; different hormones
Organ function	Kidney function (creatinine), liver function	Poor clearance → drug stays longer → more toxicity
Genetics	Gene variants affecting drug enzymes	Fast vs slow metabolisers of the same drug
Lifestyle	Smoking, alcohol, diet, exercise	Smoking speeds up breakdown of some drugs
Disease burden	Diabetes, hypertension, other illnesses	More diseases → more drugs → more interactions
Biomarkers	Blood sugar, cholesterol, gene test	Direct measures of disease state or risk

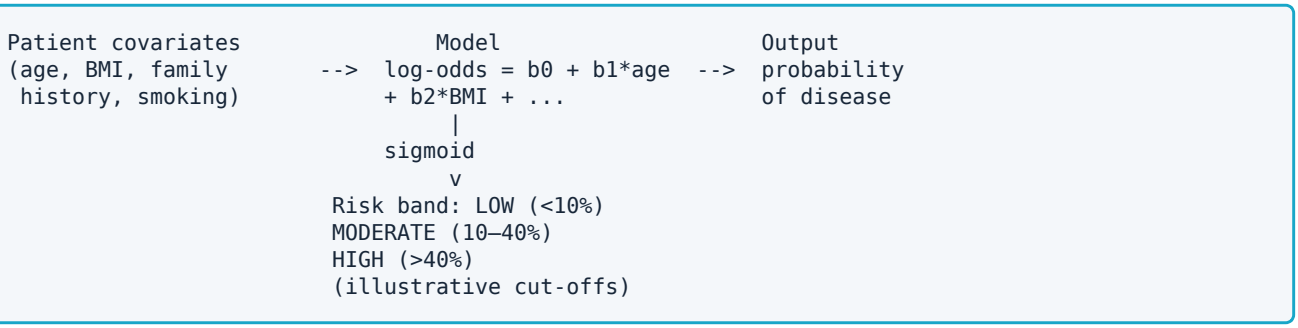
Key takeaway: A good prediction model includes covariates from several of these categories, not just one.

7.3 Logistic regression for disease risk prediction: the mechanism

Technical concept: Exactly the same mathematics as Chapter 5 (log-odds score → sigmoid → probability), but now the outcome is disease (e.g., diabetes yes/no) instead of ADR, and the predictors are risk-factor covariates (age, BMI, family history).

Simple explanation: We collect records of people with and without the disease, note their age, BMI, and other factors, and fit the S-curve. For a new person, we plug in their numbers and read off their disease probability.

Flowchart:



Key takeaway: Disease-risk prediction reuses the Chapter 5 machinery; only the question changes from "will the drug harm?" to "will the disease strike?"

7.4 WORKED EXAMPLE — predicting diabetes risk (illustrative toy dataset)

The dataset below is an **illustrative toy dataset** (invented for teaching; not real patient data, not from any real study).

Suppose a computer fitted this **illustrative** model on toy records:

log-odds (z) = −6.0 + 0.05 × age + 0.12 × BMI

Question A: What is the predicted diabetes risk for a 45-year-old with BMI 28?

Step 1 — compute the score z: z = −6.0 + 0.05 × 45 + 0.12 × 28 0.05 × 45 = 2.25 0.12 × 28 = 3.36 z = −6.0 + 2.25 + 3.36 z = −0.39

Step 2 — apply the sigmoid. e^(−z) = e^(0.39) ≈ 1.477. p = 1 ÷ (1 + 1.477) p = 1 ÷ 2.477 p ≈ 0.404

Answer: about 0.40, i.e., a 40% predicted risk — the top edge of the "moderate" band (illustrative cut-offs).

Question B: What is the predicted risk for a 30-year-old with BMI 22?

Step 1: z = −6.0 + 0.05 × 30 + 0.12 × 22 = −6.0 + 1.50 + 2.64 = −1.86 Step 2: e^(−z) = e^(1.86) ≈ 6.42 p = 1 ÷ (1 + 6.42) = 1 ÷ 7.42 ≈ 0.135

Answer: about 0.14, i.e., a 14% predicted risk — "moderate" band, much lower than Patient A.

Question C: Interpret the BMI coefficient as an odds ratio. e^0.12 ≈ 1.127. Each one-point increase in BMI multiplies the odds of diabetes by about 1.13 — roughly a 13% rise in odds per BMI point, holding age fixed (illustrative).

> **Key takeaway:** Plug the patient's numbers into the fitted equation, squeeze through the sigmoid, and you have a personalised risk number.

7.5 From risk score to action: risk bands (illustrative)

Band (illustrative cut-offs)	Predicted probability	Example action
Low	Below 0.10 (10%)	Routine advice, recheck in a few years
Moderate	0.10–0.40	Lifestyle counselling, yearly screening
High	Above 0.40	Intensive lifestyle programme, consider preventive treatment, frequent follow-up

Note: real cut-offs are chosen by clinical guidelines and cost-benefit judgement, not by the model alone. The numbers above are illustrative teaching cut-offs.

Analogy: A traffic light. Green (low) = keep going with routine care. Yellow (moderate) = slow down and watch. Red (high) = stop and act now.

7.6 Covariates and therapeutic response: why models help dosing

Technical concept: The same covariates that predict disease risk also predict drug response. Including them in models lets clinicians adjust doses for the individual instead of the average patient.

Example: Kidney function (measured by creatinine) is a covariate for many drugs cleared by the kidney. A model predicting toxicity can include creatinine as a predictor; the pharmacist then sees a higher predicted risk for a kidney patient and recommends a dose reduction.

Diagram:

```
Covariate measured        Model output          Pharmacy action
(e.g., creatinine high) → toxicity risk 45%  →  reduce dose /
                                                  extend interval /
                                                  monitor drug level
```

Key takeaway: Covariates are the bridge between "the average patient in the textbook" and "this patient in front of me".

INFOGRAPHIC CONCEPT: A tailor's shop titled "Precision Medicine: Tailor-Made Treatment". On the left, a rack of identical grey coats labeled "One-size-fits-all dosing". On the right, a tailor measuring a patient with tapes labeled Age, Weight, Genes, Kidney function, Lifestyle. In the middle, the measurements feed into a small S-curve machine that prints a ticket: "Your personal risk: 32% — recommended: lower dose + monthly check". Caption: "Measure the patient, model the risk, tailor the treatment."

7.7 Pharmacy relevance

- **Dose individualisation:** Pharmacists use covariate-based predictions to recommend dose adjustments (e.g., in kidney impairment or the elderly).
- **Screening services:** A community pharmacist can use a simple risk model (age + BMI + family history) to identify customers who should be screened for diabetes.
- **Medication review:** Risk scores help prioritise which patients need a full medication review first.

7.8 India context

- India carries a large diabetes and cardiovascular disease burden, so disease-risk models for these conditions are especially relevant to Indian pharmacy practice.
- India's population is genetically and dietarily diverse; a risk model built on one region's data should be re-validated before use in another region (recall selection bias from Chapter 6).
- Low-cost covariates (age, BMI, waist size, family history, smoking) are practical for Indian community pharmacies, where expensive gene tests are rarely available.

7.9 Key points

- Precision medicine = right drug, right dose, right patient, guided by measured covariates.
- Covariates include body size, age, sex, organ function, genes, lifestyle, diseases, and biomarkers.
- Logistic regression predicts disease risk with the same S-curve machinery used for ADR prediction.

- Risk scores are grouped into bands (low/moderate/high) that link to actions.
- Cut-offs are clinical decisions, not mathematical ones.

7.10 Common confusions

- **Precision medicine vs personalised medicine:** In this course, treat them as the same idea — tailoring care to the individual.
- **Risk factor vs cause:** High BMI raises diabetes risk (risk factor) but does not guarantee diabetes, and slim people can still get it. Models give probabilities, not certainties.
- **Covariate vs confounder:** A covariate is any measured patient characteristic we put in the model; a confounder is a covariate that specifically mixes up cause and effect (Chapter 6). Every confounder is a covariate, but not every covariate is a confounder.
- **Risk score vs diagnosis:** A 40% risk is not a diagnosis of diabetes. It means "4 out of 10 similar people would develop it" — a reason to test and prevent, not a label.

7.11 Common mistakes

- Using a risk model built on one population (e.g., one city) for a very different population without re-checking.
- Treating the risk band cut-offs as universal truths instead of illustrative, adjustable decisions.
- Forgetting lifestyle covariates (diet, smoking) and using only lab values.
- Presenting a predicted probability to a patient as a definite prediction ("you will get diabetes").

7.12 Mnemonics

- **"COVARIATE = CO-factors that VARY with the patient."**
- **"The 3 Rs of precision medicine: Right drug, Right dose, Right patient."**
- **"Green-Yellow-Red: Low-Moderate-High — the traffic light of risk bands."**

7.13 Exam summary

Precision medicine tailors treatment to individual patient covariates — body size, age, organ function, genes, lifestyle, and biomarkers. Logistic regression predicts disease risk with the same log-odds plus sigmoid machinery used for ADR prediction: patient numbers go in, a probability comes out. Risk scores are grouped into low, moderate, and high bands that guide action (routine care, counselling and screening, intensive prevention). Cut-offs are clinical choices. In India, simple low-cost covariates make such models practical for community pharmacy screening, but models must be re-validated across diverse populations.

7.14 One-minute revision

Precision = tailor treatment to covariates. Covariate examples: age, BMI, kidney function, genes, smoking. Same S-curve as Chapter 5 → disease probability. Bands: low/moderate/high → action. Probability ≠ diagnosis.

Questions

Short-answer questions:

1 Define precision medicine and covariate, with one example each.

2 Why can the same dose of a drug work differently in two patients? Name three covariates.

3 What is risk stratification? Name the three illustrative bands used in this chapter.

4 Distinguish a risk score from a diagnosis.

Long-answer questions:

1 A diabetes risk model is: log-odds = −6.0 + 0.05 × age + 0.12 × BMI. Calculate the predicted risk for (a) age 50, BMI 30 and (b) age 35, BMI 24, showing all steps. Place each patient in a risk band and suggest one action per patient.

2 Classify patient covariates into five categories with two examples each, and explain with one example how a covariate changes therapeutic response.

MCQs (answer key at the end of the chapter):

1 Precision medicine means:

a) Giving every patient the same standard dose b) Tailoring treatment to individual patient characteristics c) Using only expensive gene tests d) Treating diseases without medicines

1 Which of the following is a covariate?

a) The tablet colour b) The patient's kidney function c) The pharmacy's opening hours d) The drug's brand name

1 With log-odds = −6.0 + 0.05 × age + 0.12 × BMI, the predicted risk for age 45, BMI 28 is approximately:

a) 14% b) 27% c) 40% d) 88%

1 An odds ratio of 1.13 per BMI point means:

a) Each BMI point raises the odds of disease by about 13% b) BMI has no effect on disease c) Each BMI point lowers the odds by 13% d) 13% of patients will get the disease

1 A patient with a predicted risk of 0.08 falls in the (illustrative) band of:

a) High risk b) Moderate risk c) Low risk d) Cannot be classified

Answer key: 1-b, 2-b, 3-c, 4-a, 5-c

Viva questions:

1 What is a biomarker? Give two examples.

2 Why is kidney function an important covariate for dosing?

3 Who decides the cut-off between "moderate" and "high" risk?

4 Give one example of a protective factor.

Application-based questions:

1 As a community pharmacist, you have only a weighing scale, a BP apparatus, and a questionnaire. Design a simple diabetes-risk screening using three covariates and explain how you would counsel a "moderate risk" customer.

2 A risk model built in one Indian state is proposed for national use. List three reasons it might mislead in other states, linking each to a concept from Chapters 6–7.

References

- Aggarwal, C. C. and Reddy, C. K. Healthcare Data Analytics.
- Campbell, M. J., Machin, D. and Walters, S. J. Medical Statistics: A Textbook for the Health Sciences.
- Schmidt, S. and Derendorf, H. Applied Pharmacometrics.
- Steyerberg, E. W. Clinical Prediction Models: A Practical Approach to Development, Validation, and Updating.

CHAPTER 8: Classification of Responders vs Non-Responders; Evaluation Metrics; Ethical Implications of Predictive Modeling

Learning Objectives

After studying this chapter, you will be able to:

1 Define responder, non-responder, classifier, discrimination, calibration, decision threshold, fairness, informed consent, data privacy, and accountability.

2 Classify patients as responders or non-responders using a model probability and a threshold, with a worked toy example.

3 Explain discrimination and calibration as the two halves of model evaluation, with a worked illustration.

4 List the main ethical concerns of predictive modeling in healthcare and the pharmacist's role in each.

5 Discuss India-specific ethical and practical issues for clinical AI.

Definitions

- **Responder:** A patient in whom the treatment produces the desired effect (e.g., blood sugar falls to target, pain is relieved).
- **Non-responder:** A patient in whom the treatment does not produce the desired effect, even with correct dose and good adherence.
- **Classifier:** A model (or rule) that puts each patient into a category — here, "predicted responder" or "predicted non-responder".
- **Discrimination:** The model's ability to rank patients correctly — giving higher predicted probabilities to true responders than to true non-responders. (Nothing to do with social discrimination; it means "telling apart".)
- **Calibration:** How closely the predicted probabilities match reality. If the model says "30% chance" for 100 patients, about 30 of them should actually respond.
- **Decision threshold:** The cut-off probability that converts "predicted probability" into "predicted responder / non-responder" (often 0.5, but adjustable).
- **Fairness:** The requirement that the model works reasonably well for all patient groups, not just the majority group in the training data.
- **Informed consent:** The patient's agreement to data use or AI-assisted care after being told, in understandable words, what will happen and what the risks are.
- **Data privacy:** Protecting patients' health information from being seen, shared, or misused without permission.
- **Accountability:** Being answerable when something goes wrong — knowing who is responsible for a model's decision (the doctor? the hospital? the software maker?).
- **Algorithmic bias:** When a model systematically works worse for some group (e.g., an ethnic group or women) because the training data under-represented them.

Terminology Table (read this first)

Term	Simple meaning	Pharmacy example
Responder	Treatment worked for this patient	HbA1c reached target on metformin
Non-responder	Treatment did not work	HbA1c stayed high despite correct dose
Classifier	Model that assigns categories	Labels each patient "likely responder" or not
Discrimination	Ranks responders above non-responders	True responders mostly got higher scores
Calibration	Predicted % matches real %	"30% predicted" → about 3 in 10 responded
Decision threshold	Cut-off for the final label	0.5: ≥0.5 = predicted responder
Fairness	Works for all groups, not just the majority	Tested separately for men and women
Informed consent	Permission after clear explanation	Patient agrees to AI-assisted dose advice
Data privacy	Health data kept safe and private	No patient names leave the hospital server
Accountability	Clear responsibility for decisions	Doctor remains responsible for the final prescription

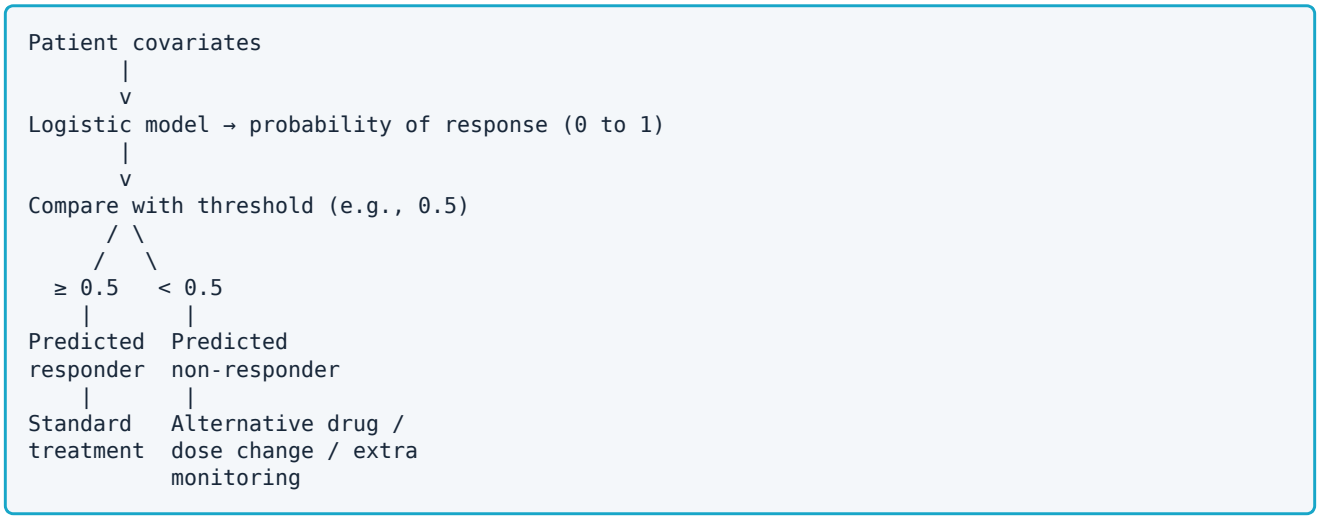
Term	Simple meaning	Pharmacy example
Algorithmic bias	Model worse for an under-represented group	Model trained mostly on men, fails for women

8.1 Classifying responders vs non-responders: the mechanism

Technical concept: A logistic regression model outputs a probability of response. Applying a decision threshold converts it into a class label: predicted responder (probability ≥ threshold) or predicted non-responder (probability < threshold). The true responder status is known only after treatment and follow-up.

Simple explanation: The model gives each patient a score out of 100. We draw a line — say at 50. Above the line: "likely responder, give the standard treatment". Below the line: "likely non-responder, consider a different drug or closer monitoring".

Flowchart:



Key takeaway: The model proposes; the threshold disposes. Choosing the threshold is a clinical decision about which error is worse (Chapter 6's trade-off applies here too).

8.2 WORKED EXAMPLE — classifying responders (illustrative toy dataset)

The dataset below is an **illustrative toy dataset** (invented for teaching; not real patient data).

A model predicts the probability that each of 12 patients will respond to a new tablet. Threshold = 0.5. After 3 months, the true response is recorded (1 = responded, 0 = did not).

Patient	Predicted prob.	Predicted class (≥0.5)	True response	Box
1	0.85	Responder	1	TP
2	0.72	Responder	1	TP
3	0.65	Responder	0	FP
4	0.58	Responder	1	TP

Patient	Predicted prob.	Predicted class (≥0.5)	True response	Box
5	0.52	Responder	0	FP
6	0.45	Non-responder	0	TN
7	0.40	Non-responder	1	FN
8	0.35	Non-responder	0	TN
9	0.30	Non-responder	0	TN
10	0.25	Non-responder	1	FN
11	0.20	Non-responder	0	TN
12	0.10	Non-responder	0	TN

Step 1 — count the boxes: TP = 3, FP = 2, FN = 2, TN = 5. Total = 12 ✓

Step 2 — metrics:

- Accuracy = (3 + 5) ÷ 12 = 8 ÷ 12 ≈ 0.667 = 66.7%
- Sensitivity = 3 ÷ (3 + 2) = 3 ÷ 5 = 0.60 = 60%
- Specificity = 5 ÷ (5 + 2) = 5 ÷ 7 ≈ 0.714 = 71.4%
- Precision = 3 ÷ (3 + 2) = 3 ÷ 5 = 0.60 = 60%

Step 3 — interpretation: The classifier correctly labels two-thirds of patients. It misses 40% of true responders (sensitivity 60%) — those patients would wrongly be denied the standard drug, so a clinician might lower the threshold or use the model only as a second opinion.

Key takeaway: Responder classification is evaluated with the same confusion-matrix tools from Chapter 6; the clinical meaning of each error decides whether the model is acceptable.

8.3 Evaluation metrics in healthcare prediction: discrimination and calibration

Technical concept: A good prediction model needs two qualities: (1) discrimination — ranking patients in the right order; (2) calibration — getting the actual probability numbers right.

Simple explanation: Discrimination is like a teacher who correctly ranks students from best to weakest, even if her marks are a bit generous. Calibration is like marks that match the true standard — a student scored "70%" actually performs at a 70% level.

Worked illustration of discrimination (illustrative toy dataset): Take 2 true responders with predicted probabilities 0.8 and 0.6, and 2 true non-responders with 0.5 and 0.3. Compare every responder with every non-responder (2 × 2 = 4 pairs):

- 0.8 > 0.5 ✓, 0.8 > 0.3 ✓, 0.6 > 0.5 ✓, 0.6 > 0.3 ✓
- 4 out of 4 pairs correctly ordered → perfect discrimination for this toy set.

In real evaluation, the fraction of correctly ordered pairs (over all responder/non-responder pairs) summarises discrimination; higher is better, 0.5 means the model ranks no better than a coin toss.

Worked illustration of calibration (illustrative toy dataset): Suppose 10 patients each got a predicted response probability near 0.30. If about 3 of the 10 actually responded, the model is well calibrated at that level — the "30%" meant what it said. If 8 of the 10 responded, the model is badly calibrated (it underestimates), even if it ranked patients correctly.

Summary table:

Quality	Question	Toy check
Discrimination	"Are responders ranked above non-responders?"	Fraction of correctly ordered pairs
Calibration	"Does 30% predicted mean ~30% actually respond?"	Compare predicted vs observed rates in groups

Key takeaway: A model can discriminate well but calibrate badly (right order, wrong numbers). Healthcare needs both: the order decides who gets priority, the numbers decide doses and honest counselling.

8.4 Classification of ethical implications

Ethical issue	Simple meaning	Pharmacy example
Fairness / algorithmic bias	Model works worse for some groups	Trained mostly on urban men; misleads for rural women
Privacy	Health data must stay protected	Patient records used for training must be de-identified and secured
Informed consent	Patients should know AI is involved	Telling a patient their dose was AI-assisted before they agree
Autonomy	The final decision stays human	Doctor/pharmacist can overrule the model; the model advises, humans decide
Accountability	Someone must answer for errors	If the model misses a fatal ADR, who is responsible — the prescriber, the hospital, or the vendor?
Transparency	Decisions should be explainable	A pharmacist should be able to explain why the model flagged a patient

The golden rule: AI advises; the healthcare professional decides. A model is a tool like a stethoscope — the clinician remains responsible for the patient.

INFOGRAPHIC CONCEPT: A courtroom-style "Ethics Checklist" poster. A judge's gavel at the top labeled "Who is accountable?". Below, six stamp boxes to tick: Fairness (balanced scales with diverse patient icons), Privacy (a locked medical file), Consent (a signed form with a pen), Autonomy (a human hand overruling a robot hand), Transparency (a magnifying glass over the model's reasoning), Safety (a shield). Bottom banner: "AI advises — the pharmacist decides." Caption: "Run every clinical model through this checklist before ward use."

8.5 Pharmacy relevance

- **Gatekeeper role:** Pharmacists sit on drug-and-therapeutics committees that approve (or reject) AI tools for the hospital — Chapters 6 and 8 give the exact questions to ask: "Show me sensitivity by subgroup. Show me calibration. Who is accountable?"
- **Patient counselling:** When a model's prediction changes therapy, the pharmacist explains it to the patient in simple words and records consent.
- **Monitoring for bias:** After deployment, pharmacists can watch whether the model under-performs in any patient group and report it like an ADR.

8.6 India context

- India's diversity (languages, diets, genetics, rural–urban divide) makes fairness testing essential: a model validated in one metro hospital may fail elsewhere.
- Health data is sensitive personal information; Indian law protects personal data, so any project using patient records (Unit V) must use de-identified data with proper permissions from the institution.
- In low-resource settings, a model that needs expensive tests is unethical in practice — it helps only the rich. Prefer models built on affordable covariates so the benefit reaches all patients.

8.7 Key points

- Responders vs non-responders are classified by applying a threshold to the predicted probability.
- Evaluation needs discrimination (correct ranking) AND calibration (honest probabilities).
- The confusion-matrix metrics from Chapter 6 apply directly to responder classification.
- Ethical issues: fairness, privacy, informed consent, autonomy, accountability, transparency.
- Golden rule: AI advises; the healthcare professional decides and remains accountable.

8.8 Common confusions

- **Discrimination (statistics) vs discrimination (society):** In modelling, "discrimination" is a compliment — it means telling responders apart from non-responders. "Algorithmic bias" is the harmful kind.
- **Calibration vs accuracy:** Accuracy counts right/wrong labels; calibration checks whether the probability numbers are honest. A model can be accurate but badly calibrated.
- **"The model decided" vs "the clinician decided":** The model only computes a number. The decision — and the responsibility — belongs to the human professional.
- **De-identified vs anonymous:** De-identified data has names/IDs removed but could theoretically be re-linked; handle it as sensitive anyway.

8.9 Common mistakes

- Deploying a model tested only on the majority group, then assuming it works for everyone.

- Using predicted probabilities for counselling without checking calibration first.
- Letting the model's label override clinical judgement ("the computer said non-responder, so no treatment").
- Collecting patient data for a project without institutional permission and consent procedures.

8.10 Mnemonics

- **"FAIR AI: Fair data, Accountable use, Informed consent, Respect privacy."**
- **"D-Cal: Discrimination ranks, Calibration counts."** (Ranking vs honest numbers.)
- **"AI Advises, Humans Decide — A-A-H-D."**

8.11 Exam summary

Patients are classified as responders or non-responders by applying a decision threshold to the model's predicted probability; performance is measured with the confusion-matrix metrics (accuracy, sensitivity, specificity, precision). Full evaluation needs discrimination (correct ranking of responders above non-responders) and calibration (predicted probabilities matching observed rates). Predictive modelling raises ethical duties: fairness across groups, data privacy, informed consent, preserving human autonomy, clear accountability, and transparency. The pharmacist's role is gatekeeper, counsellor, and bias monitor — with the golden rule that AI advises and the healthcare professional decides.

8.12 One-minute revision

Threshold on probability → responder/non-responder labels → confusion matrix metrics. Evaluation = discrimination (ranking) + calibration (honest %). Ethics: FAIR AI (fairness, accountability, informed consent, privacy) + autonomy + transparency. AI advises; humans decide.

Questions

Short-answer questions:

1 Define responder and non-responder with one pharmacy example each.

2 What is the difference between discrimination and calibration?

3 State the "golden rule" of clinical AI and explain it in one sentence.

4 What is algorithmic bias? Give one example.

Long-answer questions:

1 Using the 12-patient toy dataset in section 8.2, compute accuracy, sensitivity, specificity, and precision showing all steps, and discuss whether the classifier is acceptable for deciding who receives a costly new drug.

2 Discuss four ethical implications of predictive modelling in healthcare (fairness, privacy, informed consent, accountability), giving one pharmacy example and one safeguard for each.

MCQs (answer key at the end of the chapter):

1 A non-responder is a patient:

a) Who refuses to take medicines b) In whom the treatment does not produce the desired effect c) Who is allergic to all drugs d) Who misses clinic appointments

1 Calibration of a prediction model means:

a) Ranking responders above non-responders b) Predicted probabilities matching observed event rates c) Counting true positives d) Choosing the decision threshold

1 If 10 patients with ~30% predicted response include about 3 actual responders, the model is:

a) Well calibrated at that level b) Perfectly discriminating c) Showing algorithmic bias d) Overfitted

1 "AI advises; the healthcare professional decides" mainly protects:

a) Model accuracy b) Human autonomy and accountability c) Data storage costs d) Threshold selection

1 A model trained mostly on urban male patients that performs poorly for rural women shows:

a) Good calibration b) Algorithmic bias / unfairness c) High specificity d) Proper informed consent

Answer key: 1-b, 2-b, 3-a, 4-b, 5-b

Viva questions:

1 How do you convert a predicted probability into a responder/non-responder label?

2 Why do we need both discrimination and calibration?

3 Who is accountable when a clinical AI model makes an error?

4 Name three items on the ethics checklist for a clinical model.

Application-based questions:

1 Your hospital plans to use a responder-prediction model to decide who gets an expensive biologic drug. List five questions you would ask the vendor before approval, covering metrics and ethics.

2 A calibrated model gives your patient a 25% probability of responding to Drug A and 70% to Drug B, but Drug B costs ten times more. How would you counsel the patient, respecting autonomy and informed consent?

References

- Aggarwal, C. C. and Reddy, C. K. Healthcare Data Analytics.
- Campbell, M. J., Machin, D. and Walters, S. J. Medical Statistics: A Textbook for the Health Sciences.
- Steyerberg, E. W. Clinical Prediction Models: A Practical Approach to Development, Validation, and Updating.
- Strom, B. L., Kimmel, S. E. and Hennessy, S. Pharmacoepidemiology.