Faculty of Human and Social Sciences **Department of Psychology**

Level: MA1 Clinical Psychology **Instructor: Ms. Saci Meriem**

Module: English

Lesson n°4: Diagnostic Cutoffs Validity

Introduction

Diagnostic cutoffs (thresholds) are numeric or categorical points on measurement scales used to distinguish between cases and non-cases — for example, diseased vs. healthy, clinical vs. subclinical, or impaired vs. unimpaired. Cutoffs are central to clinical decision-making, epidemiological surveillance, screening programs, and research classification. Validity of diagnostic cutoffs determines how well these thresholds identify true cases, minimize harm from misclassification, and generalize across populations and settings. This paper examines theoretical foundations, statistical methods, threats to validity, practical considerations, and best-practice recommendations.

Conceptual foundations

I. What is a diagnostic cutoff?

A diagnostic cutoff is a rule mapping an observed score (or pattern of observations) into a binary or multi-category classification. Cutoffs may be absolute (e.g., blood glucose ≥ 126 mg/dL) or relative (e.g., top 10% on a symptom scale). They can be derived from clinical consensus, normative distributions, criterion-referenced evidence, or statistical optimization.

II. Types of validity relevant to cutoffs

- ***Criterion validity***: How well the cutoff-classification corresponds to a gold-standard diagnostic criterion (sensitivity, specificity).
- ***Construct validity***: Whether the cutoffed classification behaves as expected theoretically (relationships with risk factors, outcomes).
- ***Predictive validity:*** How well the classification predicts future outcomes (e.g., disease progression, treatment response).
- ***External validity (generalizability):*** Stability of cutoff performance across populations, settings, and measurement conditions.

III. Statistical methods for establishing cutoffs

1. Receiver Operating Characteristic (ROC) analysis

ROC curves plot sensitivity vs. 1 − specificity across thresholds. The Area Under the Curve (AUC) summarizes discriminative ability; cutoffs are often chosen to optimize a tradeoff between sensitivity and specificity (Youden's J = sensitivity + specificity − 1), to meet programmatic priorities (maximize sensitivity for screening, specificity for confirmatory tests), or to balance costs.

2. Likelihood ratios and post-test probability

Likelihood ratios (LR+ and LR−) quantify how much a test result changes odds of disease. Cutoffs with large LR+ (>10) or small LR− (<0.1) provide stronger diagnostic shifts. Bayes' theorem combines pretest probability with LRs to produce clinically useful post-test probabilities.

3. Decision curve analysis & net benefit

Decision curve analysis weights harms/benefits across thresholds and aids selection of cutoffs that maximize clinical net benefit rather than only statistical indices.

4. Classification and regression tree (CART) and machine learning approaches

When multiple predictors exist, data-driven algorithms can identify optimal decision rules that may include multiple cutoffs or complex interactions. These require careful validation to avoid overfitting.

5. Anchor- and outcome-based methods

Cutoffs can be anchored to clinically meaningful outcomes (e.g., functional impairment, hospitalization) using distribution-based methods (e.g., minimal clinically important difference) or external criteria.

IV. Evaluating cutoff validity

- ***Sensitivity (true positive rate)*** and ***Specificity (true negative rate)***.
- ***Positive Predictive Value (PPV)*** and ***Negative Predictive Value (NPV)*** — depend on prevalence.
- ***AUC (discrimination)*** — overall separability of cases/non-cases.
- ***Calibration*** — agreement between predicted probabilities (or classification) and observed outcomes.
- ***Reclassification indices*** — net reclassification improvement when comparing alternative cutoffs or models.
- ***Measurement invariance*** — whether the instrument measures the same construct across groups (essential for fair cutoffs).

V. Threats to validity

a. ***Imperfect or biased gold standards:*** If the reference standard is flawed or subjective, cutoff performance estimates will be biased (spectrum bias, incorporation bias).

b. ***Spectrum effects and case-mix variation: Performance*** depends on disease severity distribution, comorbidities, and population characteristics. A cutoff valid in tertiary care may perform poorly in primary care or community samples.

c. ***Prevalence dependence for PPV/NPV:*** Predictive values will vary with prevalence; a cutoff with high PPV in a high-prevalence clinic can have low PPV when used for population screening.

d. ***Measurement error and reliability:*** Unreliable measures widen score distributions and reduce discriminative power; optimal cutoffs shift when measurement error changes.

e. ***Cultural, linguistic, and demographic differences:*** Scores and the clinical meaning of a given score may vary across cultures, age groups, sexes, or languages, undermining generalizability.

f. ***Arbitrary or convenience-based thresholds:*** Cutoffs chosen for administrative simplicity (e.g., round numbers) may lack empirical grounding, producing misclassification.

g. ***Overfitting and optimistic estimates:*** Deriving and testing cutoffs in the same dataset inflates performance; lack of external validation risks poor real-world performance.

VI. Clinical and ethical considerations

- ***Balance of harms***: False positives may cause anxiety, unnecessary treatment, or stigma; false negatives may delay needed care. Cutoff selection should weigh relative harms.
- ***Resource implications****:* High-sensitivity cutoffs may increase downstream diagnostic burden and costs.
- ***Equity and fairness***: Cutoffs must be evaluated for differential misclassification across demographic groups to avoid exacerbating disparities.
- ***Communicating uncertainty****:* Clinicians should present risk as probability and consider cutoffs as decision aids, not absolute truths.