Application of Assessment in Clinical Settings

Definitions

- Testing
  - A scale is administered to obtain a specific score and a descriptive meaning and can be applied to the score on the basis of normative, nomothetic findings
- Assessment
  - The clinician who takes a variety of test scores, generally obtained from multiple test methods
  - Considers data in the context of history, referral information and observed behaviour to understand the person being evaluated
  - Communicate findings to the patient, significant others and referral sources
  - Develop a treatment strategy

Why

- Why assess?
  - Describe current functioning
  - Confirm, refute or modify impressions formed by clinicians
  - Identify therapeutic needs, highlight issues likely to arise in treatment, recommend forms of interventions and offer guidance about likely outcomes
  - Aid in differential diagnosis
  - Monitor treatment over time to evaluate the success of interventions
  - Manage risk
    - Untoward treatment reaction
    - Potential legal liabilities
  - Provide skilled, empathic assessment feedback as a therapeutic intervention in itself
- Why use standardised tests?
  - Clinicians are unreliable judges
    - Errors in gathering data
      - Tendency to see patterns where none exists
      - Tendency to seek confirmatory evidence
      - Use of preconceived biases
    - Error in synthesising data
      - Heuristics in clinical judgement
        - Representativeness
        - Availability
        - Anchoring
      - Affect

Type of tests

- Diagnostic interviews
  - Fully structured
    - For research or epidemiology
    - Ask question - yes or no responses
      - Move onto next question as determined by answer
  - Semi-structured
    - Initial questions can ask additional question to help with judgement
  - Ensure coverage of the diagnostic criteria as specified by DSM 5
    - Structured Clinical Interview for DSM 5
    - Few errors in gathering data
  - Rules for scoring the interview are specified
    - Few errors in synthesising data
  - Psychometric features
    - Reliability
      - Inter-rater agreement
      - Test retest reliability
    - Validity
      - Validity of diagnostic criteria
        - Diagnostic interview can only be as good as the diagnostic criteria
      - What is the "gold standard"?
        - Use a clinician not using diagnostic interview as criterion
        - LEAD standard
          - Longitudinal
          - Expert
          - All data
            - Better way to develop criterion

- Procedural validity
  - Create a 2x2 analysis

		Validation	Criterion	
		+	-	
Diagnostic Interview	+	a	b	a+b
	-	c	d	c+d
		a+c	b+d	

    - Interview twice - once diagnostic int other valid criterion
    - Say for scid for depression
      - a = valid and diag agreed diagnosis was present
      - d = valid and diag agreed diagnosis was not present
      - c = valid diagnosis was present, diag not present
      - b = valid diagnosis not present, diag was present
  - Kappa coefficient
    - Chance corrected agreement
    - $$\kappa = \frac{p_o - p_e}{1 - p_e}$$

      where:

      p_o is the observed proportion

    - $$p_o = \frac{(a+d)}{n}$$

      p_e is the proportion expected by chance

      $$p_e = \frac{(a+b)(a+c)}{n} + \frac{(c+d)(b+d)}{n}$$

    - Interpretation
      - ≥ .75 - excellent agreement
      - .6 to .74 – good agreement
      - .4 to .59 – fair agreement l
      - ess than .4 – poor agreement
      - 0 – agreement at chance level
  - About the test
    - Sensitivity
      - Probability that a person with a clinical diagnosis(validation criteria) will receive the same diagnostic interview diagnosis
      - a / a+c
      - Ability of the test to detect true positives
      - High sensitivity
        - Good at finding cases
        - Needed where cost for not finding case is high
      - May false diagnosis
    - Specificity
      - Probability that a person without a clinical diagnosis will not receive that diagnosis via the diagnostic interview
      - d / b+d
      - Ability of the test to exclude a true negative
      - High specificity
        - Good at classifying people who don't have the diagnosis
        - Needed where the cost of false positive is high

        - About the individuals
            - Positive predictive values
                - Probability that a person with a diagnostic interview is truly "ill"
                - a/a+b
                - Proportion of positive test results that are true positives
                    - Presence of disease
            - Negative predictive value
                - Probability that a person without a diagnostic interview diagnosis is truly "well"
                - d/d+c
                - Proportion of negative test results that are true negative
                    - Absence of disease
- Questionnaires
    - Delivery
        - Self report questionnaires
        - Questionnaires completed by significant others
    - Type
        - Global
            - Assess multiple symptoms
            - Provide an overall level of severity of psychopathology
            - Used for screening
            - Eg - the Brief Symptom Inventory
                - Designed to reflect the psychological symptom patterns of patients and non patients
                - Not diagnostic
                - 53 items describing psychiatric symptoms
                - Items are rated on a 5 point scale and rated on how much distress
                - Eg how much were you distressed by
                    - Nervousness
                    - Poor appetite
                    - Idea that someone else can control your thoughts
                    - Temper outbursts you could not control
                - Scored on 9 primary symptom dimensions
                - 3 global indices of distress
        - Specific
            - Short and more practical
            - Assess a limited set of symptoms
            - Provide measures of the level of severity of a specified problem
            - Used for planning treatment and monitoring progress
            - Eg - Beck Anxiety Inventory
                - 21 measure developed to assess the severity of anxiety symptoms in clinical populations
                - Aim to reliably distinguish anxiety from depression
                - Ratings of how much respondents have been bothered by each of the symptoms over the past week on a 4 points scale
                    - Unable to relax
                    - Nervous
                    - Fear of the worst happening
- Behavioural tests
    - Most commonly used in assessment of anxiety disorders
    - Eg Behavioural Avoidance tests
        - Make approach phobia
        -

Disorder	Behavioural avoidance test
Specific phobia	Client's distance from feared object
Agoraphobia	Walking distance from home
Social anxiety disorder	Delivering an impromptu speech
OCD	Touching "contaminated" objects

        - Can ask what they're feeling while its happening
            - Cognitive and physical symptoms

- Observational methods
    - Monitoring patient at home or at school

Measuring client-outcome

- An application of tests in clinical settings
- Usual method
    - Administer a test a beginning and end of treatment
    - Also give brief measures during
        - Progress monitoring
- But how do you know client is making appropriate progress or has good outcome?
    - Statistical significance
        - Different - compare group means between treatment vs no treatment
        - On average does the treatment work
            - Not about does each individual improve
    - Clinical significance of change
        - End state functioning falls within a normative range on important measures
        - Also needs to represent a reliable change
            - Needs to demonstrate
                - Improvement
                    - The dependent measure must show a reliable change that is larger than the measurement error of the instrument (Reliable Change Index)
                    - Reliable change index
                        - RC = X2 - X1/ Sdiff
                            - X1 = pretreatment score
                            - X2 = post treatment score
                            - Sdiff = standard error of the difference between the two scores
                                - Sdiff = $\sqrt{2}(SE)^2$
                        - If RC is greater than 1.96, change is reliable
                            - Don't need to know how to calculate for exam
                - Recovery
                    - After treatment, the individual's score on the dependent measure is more likely to be drawn from the distribution of a functional than a dysfunctional population
                    - Return to normal functioning
                        - 3 ways to operationalise this
                            - Post treatment score should fall outside the range of dysfunctional population, where range is 2 SDs beyond the mean
                            - Post treatment score should fall within the range of the functional population
                                - Within 2 SDs of the mean
                            - Post treatment score is closer to the mean of the functional than dysfunctional population