P686- Clinical Utility Index - A New Method to Calculate the Clinical Value of Diagnostic & Screening Tests: Proof of Concept Study

Alex J Mitchell Consultant in Psychiatry ajm80@le.ac.uk
Department of Cancer & Molecular Medicine, University Hospitals Leicester (UK)

OBJECTIVES To develop a method to help inform clinicians as to the clinical value of a diagnostic or screening tests. Conventionally such tests are evaluated using sensitivity (SE) or specificity (SP). SE and SP are simple measures of occurrence. Positive predictive value (PPV) and negative predictive value (NPV) are measures of discrimination. To be clinically useful a test should score high on occurrence and high discrimination.

METHODS A method was developed called the clinical utility index (UI). This is actually two methods. A directional measure of rule-in accuracy (also known as case-finding) using the UI+ as well as rule-out accuracy (also known as screening) using the UI-. The clinical utility index combines discriminatory ability and occurrence such that the positive utility index (UI+) = sensitivity x PPV and the negative utility index (UI-) = specificity x NPV. As proof multiple alternative measures of accuracy in a simulation model were calculated, using cancer-distress screening as an example.

Case-Finding Utility (positive utility index UI+)

= sensitivity x PPV

Screening Utility (negative utility index UI-)

= specificity x NPV

RESULTS If a hypothetical screening test [A] for cancer yielded a true positive in 9 out of 10 positive tests (90% PPV) and yet was positive in only one in five of cases (20% sensitivity) then the UI+ would be 0.18. The clinical value of test [A] would be much inferior to test [B] that was 90% accurate and positive in 90% of cases (UI+ = 0.81).

A qualitative interpretation of the clinical utility index is possible as follows **=> 0.81 is excellent; => 0.64 good; => 0.49 satisfactory/adequate; => 0.36 poor and <0.36 very poor.** When comparing diagnostic tests using the UI, it is advised to adjust for differences in prevalence. Utility index can be combined with acceptability and cost to further enhance the interpretation of test accuracy.

The following table compares various measures of accuracy from Baker-Glenn et al (2008) some of which is unpublished data against DSMIV major depression in a cancer chemotherapy setting in Leicester (UK).

Assessment of Depression	Depressed	Identified	Sensitivity	Non-Depressed	Identified	Specificity	PPV	NPV	UI+	UI-	Fraction Correct
PHQ2 (depression)	24	16	66.7%	191	188	98.4%	84.2%	95.9%	56.1%	94.4%	94.9
PHQ1 (interest)	24	21	87.5%	191	179	93.7%	63.6%	98.4%	55.7%	92.2%	93.0
PHQ9 (11v12)	24	23	95.8%	191	185	96.9%	79.3%	99.5%	76.0%	96.3%	96.7
PHQ9 (9v10)	24	24	100.0%	191	166	86.9%	49.0%	100.0%	49.0%	86.9%	88.4
PHQ8	24	24	100.0%	191	157	82.2%	41.4%	100.0%	41.4%	82.2%	84.2
DT (4v5)	24	16	66.7%	193	152	78.8%	28.1%	95.0%	18.7%	74.8%	77.4
DT (3v4)	24	19	79.2%	193	138	71.5%	25.7%	96.5%	20.3%	69.0%	72.4
DepT (5v6)	24	14	58.3%	193	180	93.3%	51.9%	94.7%	30.2%	88.4%	89.4
DepT (6v7)	24	10	41.7%	193	188	97.4%	66.7%	93.1%	27.8%	90.7%	91.2
HADS-A	24	19	79.2%	193	127	65.8%	22.4%	96.2%	17.7%	63.3%	67.3
HADS-D	24	19	79.2%	193	169	87.6%	44.2%	97.1%	35.0%	85.0%	86.6
HADS-T	24	23	95.8%	193	154	79.8%	37.1%	99.4%	35.6%	79.3%	81.6
Help QQ Alone	24	10	41.7%	193	168	87.0%	28.6%	92.3%	11.9%	80.4%	82.0
Help QQ AND PHQ2	24	10	41.7%	193	188	97.4%	66.7%	93.1%	27.8%	90.7%	91.2

CONCLUSIONS The interpretation of diagnostic and screening tests can be improved by combining occurrence with discrimination. The clinical utility index offers this with a simple qualitative and quantitative interpretation.