Levels of Evidence

What are the levels of evidence?

These decisions gives the "grade (or strength) of recommendation." **Evidence** from a systematic review or meta-analysis of all relevant RCTs (randomized controlled trial) or **evidence**-based clinical practice guidelines based on systematic reviews of RCTs or three or more RCTs of good quality that have similar results. Feb 2, 2017

Levels of evidence (sometimes called hierarchy of evidence) are assigned to studies based on the methodological quality of their design, validity, and applicability to patient care. These decisions gives the "grade (or strength) of recommendation."

Level of evidence (LOE)	Description
Level I	Evidence from a systematic review or meta-analysis of all relevant RCTs (randomized controlled trial) or evidence-based clinical practice guidelines based on systematic reviews of RCTs or three or more RCTs of good quality that have similar results.
Level II	Evidence obtained from at least one well-designed RCT (e.g. large multi-site RCT).
Level III	Evidence obtained from well-designed controlled trials without randomization (i.e. quasi-experimental).
Level IV	Evidence from well-designed case-control or cohort studies.
Level V	Evidence from systematic reviews of descriptive and qualitative studies (meta-synthesis).
Level VI	Evidence from a single descriptive or qualitative study.
Level VII	Evidence from the opinion of authorities and/or reports of expert committees.

This level of effectiveness rating scheme is based on the following: Ackley, B. J., Swan, B. A., Ladwig, G., & Tucker, S. (2008). *Evidence-based nursing care guidelines: Medical-surgical interventions.* (p. 7). St. Louis, MO: Mosby Elsevier.

From the Centre for Evidence-Based Medicine, Oxford

For the most up-to-date levels of evidence, see www.cebm.net/?o=1025

Therapy/Prevention/Etiology/Harm:

1a:	Systematic reviews (with homogeneity) of randomized controlled trials
1b:	Individual randomized controlled trials (with narrow confidence interval)
1c:	All or none randomized controlled trials
2a:	Systematic reviews (with homogeneity) of cohort studies
2b:	Individual cohort study or low quality randomized controlled trials (e.g. <80% follow-up)
2c:	"Outcomes" Research; ecological studies
3a:	Systematic review (with homogeneity) of case-control studies
3b:	Individual case-control study
4:	Case-series (and poor quality cohort and case-control studies)
5:	Expert opinion without explicit critical appraisal, or based on physiology, bench research or "first principles"

From the Centre for Evidence-Based Medicine, Oxford

Continued

Diagnosis:

1a:	Systematic review (with homogeneity) of Level 1 diagnostic studies; or a clinical decision rule with 1b studies from different clinical centers.
1b:	Validating cohort study with good reference standards; or clinical decision rule tested within one clinical center
1c:	Absolute SpPins And SnNouts (An Absolute SpPin is a diagnostic finding whose Specificity is so high that a Positive result rules-in the diagnosis. An Absolute SnNout is a diagnostic finding whose Sensitivity is so high that a Negative result rules-out the diagnosis).
2a:	Systematic review (with homogeneity) of Level >2 diagnostic studies
2b:	Exploratory cohort study with good reference standards; clinical decision rule after derivation, or validated only on split-sample or databases
3a:	Systematic review (with homogeneity) of 3b and better studies
3b:	Non-consecutive study; or without consistently applied reference standards
4:	Case-control study, poor or non-independent reference standard
5:	Expert opinion without explicit critical appraisal, or based on physiology, bench research or "first principles"

From the Centre for Evidence-Based Medicine, Oxford

Continued

Prognosis:

1a:	Systematic review (with homogeneity) of inception cohort studies; or a clinical decision rule validated in different populations.
1b:	Individual inception cohort study with > 80% follow-up; or a clinical decision rule validated on a single population
1c:	All or none case-series
2a:	Systematic review (with homogeneity) of either retrospective cohort studies or untreated control groups in randomized controlled trials.
2b:	Retrospective cohort study or follow-up of untreated control patients in a randomized controlled trial; or derivation of a clinical decision rule or validated on split-sample only
2c:	"Outcomes" research
4:	Case-series (and poor quality prognostic cohort studies)
5:	Expert opinion without explicit critical appraisal, or based on physiology, bench research or "first principles"

Note: A minus sign "-" may be added to denote evidence that fails to provide a conclusive answer because it is *either* (a) a single result with a wide Confidence Interval; *OR* (b) a Systematic Review with troublesome heterogeneity. Such evidence is inconclusive, and therefore can only generate Grade D recommendations.

Strength-of-Recommendation Taxonomy (SORT)

Code	Definition
A	Consistent, good-quality patient-oriented evidence *
B	Inconsistent or limited-quality patient-oriented evidence *
C	Consensus, disease-oriented evidence *, usual practice, expert opinion, or case series for studies of diagnosis, treatment, prevention, or screening

* Patient-oriented evidence measures outcomes that matter to patients: morbidity, mortality, symptom improvement, cost reduction, and quality of life. Disease-oriented evidence measures immediate, physiologic, or surrogate end points that may or may not reflect improvements in patient outcomes (e.g. blood pressure, blood chemistry, physiologic function, pathologic findings).

Grading of Recommendations Assessment, Development and Evaluation (GRADE)

Code	Quality of Evidence	Definition
A	High	Further research is very unlikely to change our confidence in the estimate of effect. • Several high-quality studies with consistent results • In special cases: one large, high-quality multi-centre trial
B	Moderate	Further research is likely to have an important impact on our confidence in the estimate of effect and may change the estimate. • One high-quality study • Several studies with some limitations
C	Low	Further research is very likely to have an important impact on our confidence in the estimate of effect and is likely to change the estimate. • One or more studies with severe limitations
D	Very Low	Any estimate of effect is very uncertain. • Expert opinion • No direct research evidence • One or more studies with very severe limitations

Source: GRADE (Grading of Recommendations Assessment, Development and Evaluation) Working Group 2007 1 (modified by the EBM Guidelines Editorial Team)

Key to Interpretation of Practice Guidelines

Agency for Healthcare Research and Quality:

A:	There is good research-based evidence to support the recommendation.
B:	There is fair research-based evidence to support the recommendation.
C:	The recommendation is based on expert opinion and panel consensus.
X:	There is evidence of harm from this intervention.

USPSTF Guide to Clinical Preventive Services:

A:	There is good evidence to support the recommendation that the condition be specifically considered in a periodic health examination.
B:	There is fair evidence to support the recommendation that the condition be specifically considered in a periodic health examination.
C:	There is insufficient evidence to recommend for or against the inclusion of the condition in a periodic health examination, but recommendations may be made on other grounds.
D:	There is fair evidence to support the recommendation that the condition be excluded from consideration in a periodic health examination.
E:	There is good evidence to support the recommendation that the condition be excluded from consideration in a periodic health examination.

University of Michigan Practice Guideline:

A:	Randomized controlled trials.
B:	Controlled trials, no randomization.
C:	Observational trials.
D:	Opinion of the expert panel.

Key to interpretation of practice guidelines

Continued

Other guidelines:

A:	There is good research-based evidence to support the recommendation.
B:	There is fair research-based evidence to support the recommendation.
C:	The recommendation is based on expert opinion and panel consensus.
X:	There is evidence that the intervention is harmful.