

Study type	Question	Level I	Level II	Level III	Level IV	Level V
Diagnostic—Investigating a diagnostic test	Is this (early detection) test worthwhile?	<ul style="list-style-type: none"> Randomized controlled trial 	<ul style="list-style-type: none"> Prospective^[3] cohort^[4] study 	<ul style="list-style-type: none"> Retrospective^[5] cohort^[4] study Case-control^[6] study 	<ul style="list-style-type: none"> Case series 	<ul style="list-style-type: none"> Mechanism-based reasoning
	Is this diagnostic or monitoring test accurate?	<ul style="list-style-type: none"> Testing of previously developed diagnostic criteria (consecutive patients with consistently applied reference standard and blinding) 	<ul style="list-style-type: none"> Development of diagnostic criteria (consecutive patients with consistently applied reference standard and blinding) 	<ul style="list-style-type: none"> Nonconsecutive patients No consistently applied reference standard 	<ul style="list-style-type: none"> Poor or nonindependent reference standard 	<ul style="list-style-type: none"> Mechanism-based reasoning
Prognostic—Investigating the effect of a patient characteristic on the outcome of a disease	What is the natural history of the condition?	<ul style="list-style-type: none"> Inception^[3] cohort study (all patients enrolled at an early, uniform point in the course of their disease) 	<ul style="list-style-type: none"> Prospective^[3] cohort^[4] study (patients enrolled at different points in their disease) Control arm of randomized trial 	<ul style="list-style-type: none"> Retrospective^[5] cohort^[4] study Case-control^[6] study 	<ul style="list-style-type: none"> Case series 	<ul style="list-style-type: none"> Mechanism-based reasoning
Therapeutic—Investigating the results of a treatment	Does this treatment help? What are the harms? ^[7]	<ul style="list-style-type: none"> Randomized controlled trial 	<ul style="list-style-type: none"> Prospective^[3] cohort^[4] study Observational study with dramatic effect 	<ul style="list-style-type: none"> Retrospective^[5] cohort^[4] study Case-control^[6] study 	<ul style="list-style-type: none"> Case series Historically controlled study 	<ul style="list-style-type: none"> Mechanism-based reasoning
Economic	Does the intervention offer good value for dollars spent?	Computer simulation model (Monte Carlo simulation, Markov model) with inputs derived from Level-I studies, lifetime time duration, outcomes expressed in dollars per quality-adjusted life years (QALYs) and uncertainty examined using probabilistic sensitivity analyses	Computer simulation model (Monte Carlo simulation, Markov model) with inputs derived from Level-II studies, lifetime time duration, outcomes expressed in dollars per QALYs and uncertainty examined using probabilistic sensitivity analyse	Computer simulation model (Markov model) with inputs derived from Level-II studies, relevant time horizon, less than lifetime, outcomes expressed in dollars per QALYs and stochastic multilevel sensitivity analyses	Decision tree over the short time horizon with input data from original Level-II and III studies and uncertainty is examined by univariate sensitivity analyses	Decision tree over the short time horizon with input data informed by prior economic evaluation and uncertainty is examined by univariate sensitivity analyses

^[1]This chart was adapted from "OCEBM Levels of Evidence Working Group" * and "The Journal of Bone&Joint Surgery" **. A glossary of terms can be found here: <http://www.cebm.net/glossary/>.

*The Oxford 2011 Levels of Evidence, Oxford Centre for Evidence-Based Medicine, <http://www.cebm.net/ocbm-levels-of-evidence/>.

**<http://bjbs.org/instructions-for-authors#LevelofEvidence>.

^[2] Level-I through IV studies may be graded downward on the basis of study quality, imprecision, indirectness, or inconsistency between studies or because the effect size is very small; these studies may be graded upward if there is a dramatic effect size. For example, a high-quality randomized controlled trial (RCT) should have ≥80% follow-up, blinding, and proper randomization. The Level of Evidence assigned to systematic reviews reflects the ranking of studies included in the review (i.e., a systematic review of Level-II studies is Level II). A complete assessment of the quality of individual studies requires critical appraisal of all aspects of study design.

^[3] Investigators formulated the study question before the first patient was enrolled.

^[4] In these studies, "cohort" refers to a nonrandomized comparative study. For therapeutic studies, patients treated one way (e.g., cemented hip prosthesis) are compared with those treated differently (e.g., cementless hip prosthesis).

^[5] Investigators formulated the study question after the first patient was enrolled.

^[6] Patients identified for the study on the basis of their outcome (e.g., failed total hip arthroplasty), called "cases," are compared with those who did not have the outcome (e.g., successful total hip arthroplasty), called "controls."

^[7] Sufficient numbers are required to rule out a common harm (affects >20% of participants). For long-term harms, follow-up duration must be sufficient.