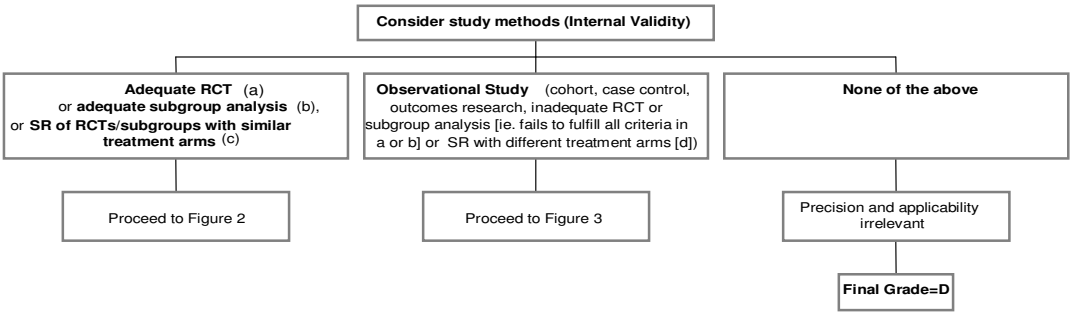
Figure 1: Algorithm for assigning evidence grades to therapy recommendations



Definitions:

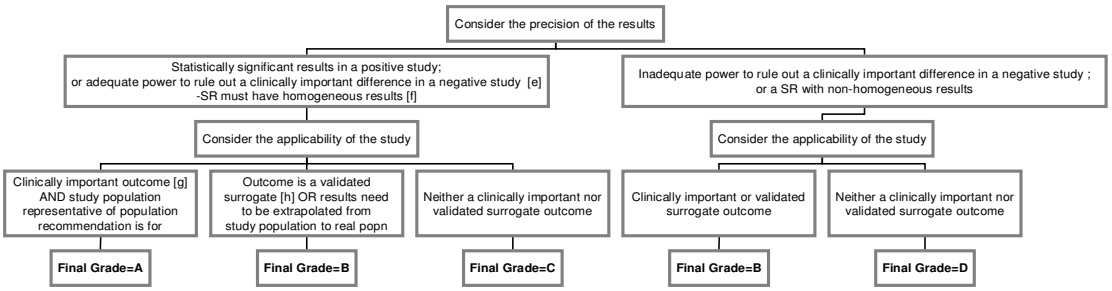
a Randomized clinical trial with blinded assessment of outcomes (if applicable), intention-to-treat analysis, adequate follow-up (ie. at least 90%, or losses to follow-up are too few to materially affect the results), and sufficient sample size to detect a clinically important difference with power > 80%.

b Subgroup analysis was a-priori, done within an adequate RCT, one of only a few tested, and there was sufficient sample size within the examined subgroup to detect a clinically important difference with power > 80%.

c Sytematic review (SR, also known as meta-analysis) in which the comparison arms are derived from head-to-head comparisons within the same RCT.

d SR in which the comparison arms are derived from different placebo-controlled RCTs, then extrapolations are made across RCTs.

Figure 2: Algorithm for assigning evidence grades to therapy recommendations (continued from figure 1- for adequate randomized trials, systematic reviews, or subgroup analyses)



Definitions:

e Adequate power in a negative study implies that 95% CI exclude a clinically important difference.

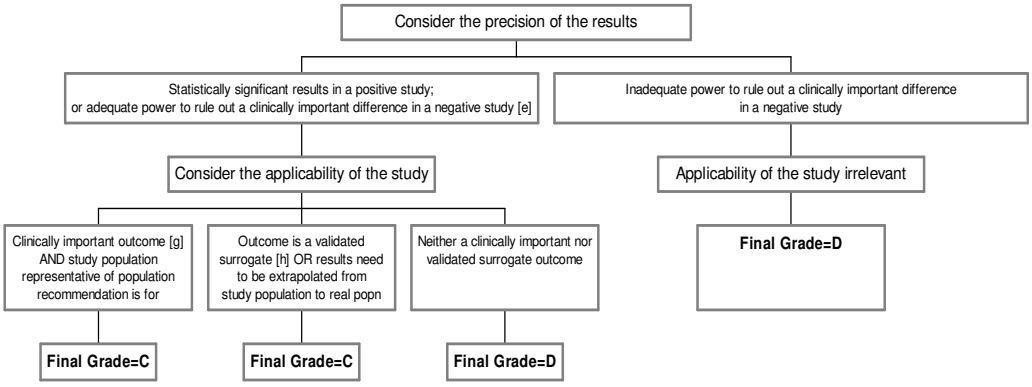
f Effect estimates in each study included in the systematic review are qualitatively similar (ie. in the same direction).

g "Hard" endpoints such as death, stroke, myocardial infarction, hospitalization, and need for dialysis; or measures of quality of life.

h Endpoints which have been consistently shown to be associated with the clinical end point in multiple studies (observational or RCT), and RCTs have consistently demonstrated that improvement in the surrogate translates into a consistent and predictable improvement in the clinical end point.

Appendix to: Levin A, Hemmelgarn B, Culleton B, et al. Guidelines for the management of chronic kidney disease. *CMAJ* 2008;179:1154-62.

Figure 3: Algorithm for assigning evidence grades to therapy recommendations
(continued from Figure 1 - for observational studies)



Definitions:

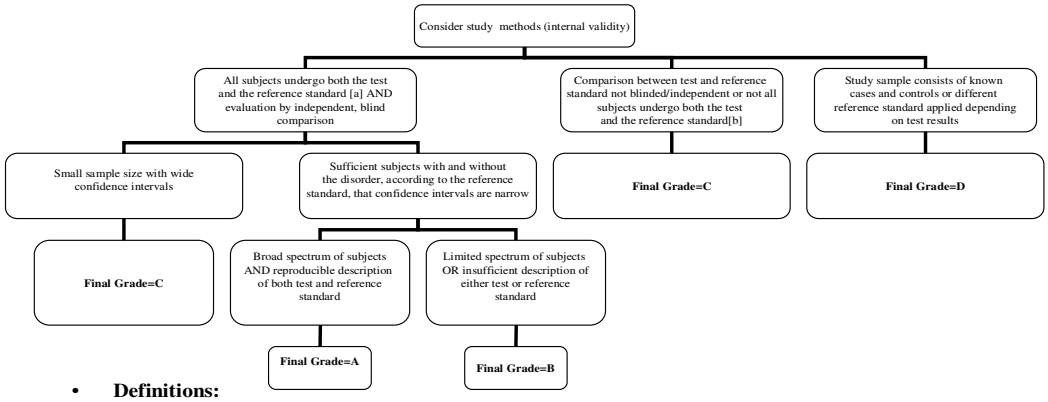
e Adequate power in a negative study implies that 95% CI exclude a clinically important difference.

f Effect estimates in each study included in the systematic review are qualitatively similar (ie. in the same direction).

g "Hard" endpoints such as death, stroke, myocardial infarction, hospitalization and need for dialysis; or measures of quality of life.

h Endpoints which have been consistently shown to be associated with the clinical end point in multiple studies (observational or RCT), and RCTs have consistently demonstrated that improvement in the surrogate translates into a consistent and predictable improvement in the clinical end point.

Figure 4: Algorithm for assigning evidence grades to diagnostic recommendations



- **Definitions:**

a The gold standard. This can be either another test which is currently accepted as the gold standard or analysis of a representative cohort of patients who underwent the test of interest and are followed for a sufficient length of time that occurrence of the target outcome is likely if the diagnosis is present (with adjustment for covariates associated with prognosis).

b Note that if follow-up of a cohort is not sufficiently long or complete enough to rule out diagnostic errors, or if data is not adjusted for covariates, then this category would apply.

Appendix to: Levin A, Hemmelgarn B, Culleton B, et al. Guidelines for the management of chronic kidney disease. *CMAJ* 2008;179:1154-62.