GRADING OF EVIDENCE FOR PROGNOSTIC QUESTIONS
American College of Emergency Physicians Clinical Policies Committee

Citation: _____

Design 1	Prospective cohort study
Design 2	Retrospective cohort study; Cross-sectional study; Case-control study
Design 3	Case series

Applicable to Clinical Question: Direct Indirect Not

Dimensions for Grading (consider both quality of execution and importance to result):
NR/NA/U: Not reported not applicable, or unclear.

			Comments
Inclusion criteria defined & appropriate	Y	N	NR / NA / U
Appropriate sampling	Y	N	NR / NA / U
Studied at a uniform time in their disease course	Y	N	NR / NA / U
Risk factor measured in a valid and reliable way	Y	N	NR / NA / U
Risk factor measured without knowledge of outcome	Y	N	NR / NA / U
Outcome measured without knowledge of risk factor	Y	N	NR / NA / U
Outcome measured in a valid and reliable way	Y	N	NR / NA / U
Appropriate level of attrition	Y	N	NR / NA / U
Accounting for drop-outs	Y	N	NR / NA / U
Appropriate sample size	Y	N	NR / NA / U
Generalizability	Y	N	NR / NA / U
Data managed appropriately	Y	N	NR / NA / U
Analyses appropriate	Y	N	NR / NA / U
Conclusions supported by the results	Y	N	NR / NA / U
Industry sponsored	Y	N	NR / NA / U

Downgrading:
- ☐ No downgrading (no methodological limitations and directly applicable)
- ☐ Downgrade 1 level (only minor methodological limitations)
- ☐ Downgrade 1 level (indirectly applicable)
- ☐ Downgrade 2 levels (major methodological limitation[s])
- ☐ Fatally flawed or not applicable

Class of Evidence: **I** **II** **III** **X**

Notes:

Reviewer: _____ **Date:** _____

Guidelines for Use:

1. Use the top grid to assign a **Design** (1, 2 or 3) based on the study's design. Some designs may not fit this schema and should be assessed individually.
2. To qualify as **Design 1**, a study of prognosis should include patients with and without the risk factor of interest who are followed prospectively and evaluated to determine if they develop the outcome of interest. Some prognostic studies are done backwards; instead of finding patients with and without a risk factor and following them to see if they develop the outcome, case-control studies select patients who have developed or not developed the outcome and look retrospectively to see if the risk factor was present. Case-control studies are designated **Design 2**. Cross-sectional studies, where the presence of the risk factor and outcome are determined simultaneously, should also be considered **Design 2**.
3. **Applicability to the clinical question** relates to whether the study being evaluated is directly, indirectly, or not applicable to the clinical question proposed as part of the clinical policy.
4. Then assess the quality of the execution of the study using the list of important dimensions as reminders. Important dimensions to be considered when assessing the quality of a study include:
   a. A clear description of how patients were included in the study, including explicit and appropriate **inclusion and exclusion criteria** and appropriate **sampling** to generate the study sample from the base population. If conducted poorly, one or both features may introduce selection bias.
   b. The **presence of the risk factor** should be determined at a uniform time early in the course of the disease.
   c. The presence or absence of a **risk factor** should be determined in a valid and reliable fashion.
   d. The presence or absence of the **outcome** of interest should be determined in a valid and reliable fashion.
   e. To avoid **expectation and recall bias**, the presence of risk factor should be determined without knowledge of the outcome and vice versa. Prospective cohort studies automatically fulfill this criterion. This criterion can be disregarded if the risk factor is an objective measure (e.g., gender). Also, the outcome of interest should ideally be measured without knowledge of the presence of the risk factor. This criterion can be disregarded if the outcome measure is objective (e.g., mortality).
   f. **Attrition** (i.e., patients who do not complete the study) may significantly bias a study. The impact of patients dropping out of a study can be lessened by various statistical techniques (e.g., intention-to-treat, imputation, sensitivity analyses). In general, attrition should be <20% but in some instances should be significantly less.
   g. **Sample size** should be sufficient to provide adequate precision of estimates and to prevent type II errors (i.e., not finding a difference when one actually exists).
   h. **Generalizability** refers to the ability to generalize the study's results to other patients or settings. Consider the representativeness of the patient population included in the study (e.g., were only patients with severe disease included).
   i. **Data management** refers to whether the data were appropriately handled during collection and analyses; this may include issues of use of a data safety monitoring board, whether authors had access to data, and who performed analyses.
   j. **Analyses** should be appropriate and valid for the study design (e.g., appropriate use of multivariable methods, confounders, interactions, collinearity, model fit; use of propensity methods; appropriate handling of missing data; etc.). k. **Conclusions supported by results** refers specifically to whether the conclusions are appropriately aligned with reported results or whether the authors took liberty in over- or under-extending their conclusions.
   l. **Industry sponsored** studies often are influenced, either in their design, performance, or reporting, by the company, which may introduce bias. Who controlled and analyzed the data? Was a DSMB used?
5. At the **Downgrading** section, summarize the quality of execution and applicability to the clinical question into a decision on downgrading. The idea here is that the maximum evidence class that can be assigned is limited by the Design (i.e., Design 1 can support up to Class of Evidence I, but Design 2 can only support Class of Evidence II or lower, and so on). Essentially, the quality of execution is used to "downgrade" studies from the maximum class, as shown in the table below. Additionally, applicability to the clinical question also relates to downgrading. (e.g., not applicable studies receive a Class of Evidence "X"). Evidence Class X studies will not be used to support clinical policies. Use the downgrading results to generate a **Class of Evidence** based on the table below.

	Design		
Downgrading	1	2	3
None	I	II	III
1 level	II	III	X
2 levels	III	X	X
Fatally flawed or NA	X	X	X