Delfini Appraisal Tool
Registry Evaluation Checklist: Safety of Interventions

KEY QUESTION

Are the results likely to be true? Or are there biases which make the intervention appear safer than it is? Or which make it appear to cause greater harm than it does, but which are actually caused by something other than the intervention under study?

1. Selection Biases
    a. Is there actual or potential bias concerning registry participants? Examples: participants are not new starts (inception cohort); health care access issues; potential for missed capture; voluntary submission as compared to mandatory; if mandatory, likelihood of complete capture; retrospective capture; incorrect diagnosis; potential for double-counting; healthy user effects affecting population for study if patients are no longer treated with agent due to adverse effects, thereby diminishing the at risk population over time.
    b. Could exclusions make the intervention seem safer than it actually is?
    c. Are there likely biases with the comparison group(s)? Examples: not including group with similar disease spectrum and not treated with agent of interest. While comparison to the general population may be useful, be mindful that this may be biased (example, different care experiences, different opportunities for cancer detection, etc); comparator groups that are not concurrent.
    d. Are there key differences between the group of interest and the comparison groups that could distort outcomes? Evaluate choice of characteristics for adjustments made between groups. Example: ethnicity, genetic issues, etc.
    e. Could discontinuations make the treatment appear safer or result in a population with a different risk profile?
    f. Are findings from the populations generalizable to the larger population?
2. Performance Biases
    a. Are there potential biases in how patients are managed (e.g., subcutaneous versus IV administration of drugs) and followed (e.g., frequency of visits) that could affect the outcomes?
    b. Are there time-based issues that could distort results such as changes in exposure to drugs, care experiences or detection of outcomes over time?
3. Data Collection Biases
    a. Are there actual or likely biases in data accuracy? Consider the impact of losses and withdrawals.
    b. Was the frequency of data collection described and is it appropriate?
    c. Is the temporal relationship appropriate for determining potential causality?
    d. Is data capture incomplete or likely to be incomplete in significant ways? Examples: comorbidities; co-interventions including dosing, duration and adherence; risk factors; protective or mitigating effects?
    e. Were switches to other interventions captured?
    f. Were quality control checks performed?
4. Assessment Biases
    a. Were outcomes of interest defined *a priori*?
    b. Were assessors blinded?
    c. Are outcomes clinically significant? (Review confidence intervals.)
    d. Are definitions for things like treatment success or failure, adverse events, etc., reasonable?
    e. Are composite endpoints appropriate?
    f. Is the risk window appropriate?
    g. Do variations in findings raise questions of potential inconsistency that may indicate the operation of bias? Example: Findings of higher rates of outcomes in earlier time periods may be reflective of a change to a population at lesser risk (healthy user effect) if patients at risk diminish from the pool over time.
    h. Is the safety population limited to only those patients who received the intervention?
    i. Were switches to other interventions handled appropriately?
    j. Could missing values make the intervention appear safer than it actually is?
    k. Were missing values handled appropriately?
    l. Were analysis methods appropriate?
    m. If modeling was used, were the assumptions used appropriate?
    n. Is there a potential for bias if assessment is not blinded?
    o. Might non-significance be explained by power issues?
    p. Were appropriate and useful sensitivity analyses performed?
    q. Is reporting inappropriately selective?

See also Delfini 1-Pager on **Safety**.

© Delfini Group, LLC, 2010. All Rights Reserved Worldwide.

Suggested Stock Language to Describe Potential Bias in Registry Studies as Applicable

- Differences might be reflective of detection and not causation.
- Patients on the agents of interest may have been more or less at risk for the development of outcome of interest due to channeling.
- Patients on the agents of interest may have been more at risk for overdiagnosis bias if there are pretreatment, care or evaluation differences between groups.
- Potential for differences in baseline characteristics in the groups.
- Lack of blinding could distort the results.
- Potential for different co-interventions, clinical management, measurement and assessment differences may confound results.
- More follow-up time may be needed.
- Summary: High Risk of Bias—Potential for many biases including different prognostic variables in the groups, detection methods, clinical management, co-interventions, measurement and assessment differences, which may confound results.

© Delfini Group, LLC, 2010. All Rights Reserved Worldwide.