Suggestions for Readers of Research Using Statistics

Overarching Guidelines:

1. Look for sources of uncertainty.
2. Keep these quotes in mind:
   "Some experts think peer review validates published research. For those of us who have been editors, associate editors, reviewers, or the targets of peer review, this argument may ring hollow. Even for careful readers of journal articles, the argument may seem a little farfetched."
      David A. Freedman, Chance 2008, v. 21 No. 1, p. 61

   "... scientists understand that peer review per se provides only a minimal assurance of quality, and that the public conception of peer review as a stamp of authentication is far from the truth.
      Charles G. Jennings, Quality and Value: The true purpose of peer review?, Peer-To-Peer, Nature.com 2006

Specific Suggestions:

- Do not just read the abstract.
  - Abstracts sometimes focus on conclusions that are more speculative than the data warrant.
- Identify the exact research question(s) the researchers are asking
  - Decide if this is a questions that you are interested in.
- Identify the measures the researchers are using. For example, if you are interested in the effect of a medication on the incidence of hip fractures, is this the endpoint that the researchers have studied, or have they just studied a proxy such as bone density?
- Determine the type of study: observational or experimental; exploratory of confirmatory
  - This will influence the strength of the conclusions that can be drawn; generally speaking, experimental studies give stronger evidence than observational studies, and confirmatory studies give stronger evidence than exploratory studies.
- Pay attention to how the sample(s) was/were chosen.
  - Think about any circumstances that might make the sample biased. Remember that results from a biased sample are unreliable, although sometimes they might give some information about a smaller population than intended.
  - Remember that voluntary response samples are usually biased.
- Have the researchers explained why the statistical procedures they have used are appropriate for the data they are analyzing?
  - In particular, have they given good reasons why the model assumptions fit the context well enough?

- o If not, their results should be given less credibility than if the model has been shown to fit the context well.
- If there is multiple inference using the same data, have the authors taken that into account appropriately in deciding significance or confidence levels?
- If hypothesis tests are used, are confidence intervals also given?
  - o Confidence intervals can give an idea of the range of uncertainty due to sampling variability.
  - o But be aware that there might also be other sources of uncertainty not captured by confidence intervals (e.g., bias, lack of fit of model assumptions, measurement uncertainty.)
- Have claims been limited to the population from which the data were actually gathered?
- Have the authors taken practical significance as well as statistical significance into account in drawing conclusions?
- Is the power of statistical tests large enough to warrant claims of no difference?
- See Good and Hardin (2010, Chapter 9) for more suggestions and details.
- See van Belle (2008, Chapter 7) for items specific to Evidence Based Medicine

Further References:

Good, P. I. and Hardin, J. W, (2010). *Common Errors in Statistics*, 3rd ed., Wiley

van Belle, Gerald (2008). Statistical Rules of Thumb, 2nd ed., Wiley

Note: For an online version of this with embedded links, see
http://www.ma.utexas.edu/users/mks/statmistakes/readers.html