EVALUATING CLINICAL LITERATURE: Part three of a four part series on study design, statistics, interpreting results and bias.

Interpreting Confidence Intervals

- Confidence intervals can help you determine whether study results are weak or strong, definitive or not definitive. They are also more informative than a single value such as the p-value when assessing the strength of the evidence. To put it another way, confidence intervals are similar to a margin of error.

- Confidence intervals (CI) are an estimated range of values that attempt to quantify uncertainty. A narrow confidence interval implies high precision, whereas a wide interval implies poor precision, which is often an indication of an inadequate sample size. Most studies report CI as 95%, which means that if you conducted 100 identical studies, 95 of them would have results within the 95% CI. If the 95% CI includes zero (i.e., -1.2 to 2.4) then the outcome is NOT considered statistically significant because we cannot rule out the possibility that there is no effect. For example, a study of homeopathic treatment of pain and swelling after oral surgery (Lokken, 1995) examined swelling three days after an operation. The study showed that homeopathy led to a 1mm less swelling on average. The 95% confidence interval ranged from –5.5 to 7.5 mm, which appears to be a wide interval and implies that neither a large improvement due to homeopathy nor a large decrement could be ruled out.

- When reviewing a published medical report and interpreting confidence intervals, you should look for two things:
  - Does the interval contain a value that implies no change or no effect? Does the interval include zero?
  - Does the confidence interval lie partly or entirely within a range of clinical indifference?
    - Clinical indifference represents values of such a trivial size that you would not want to change your current practice. For example, you would not prescribe a statin that lowered LDL by 1%.

- When looking at a negative trial, look to the upper end of the confidence interval, meaning the end that suggests the largest benefit from treatment. If even the smallest benefit of clinical importance lies above the upper limit of the confidence interval, the trial is definitively negative. In contrast, if clinically important benefits fall within the confidence interval, the trial has not ruled out the possibility that the treatment is worth while.

Reporting Results

Absolute Risk Reduction (ARR)- This is the absolute difference in outcome rates between the control and treatment groups. If 6% of the control group dies and 3% of the treatment group dies, the ARR is 3%.

Relative Risk Reduction (RRR)- Relative risk measures how much the risk is reduced in the experimental group compared to a control group. In our example above, the treatment would have a relative risk reduction of .5 or 50%. (The rate of death in the treated group is half of that in the control group.) RRR does not reflect the true magnitude of the treatment effect and its use can make treatment seem more effective (as in our example above), but it can also make adverse events appear more alarming. Be skeptical of any study or news release that reports the relative risk reduction without also reporting the absolute risk reduction.

Odds Ratio (OR)- Rather than looking at the risk of an event, the odds ratio looks at the odds of the event occurring. An odds ratio of 1 means that there is no difference in the odds of the event occurring between the groups. The further the OR is from 1, the greater the difference in odds between the two groups. The odds ratio is useful in retrospective studies such as case-control studies, where absolute risk and relative risk cannot be used (these measures can only be used in prospective studies). Odds ratios are comparable to relative risk when the outcomes being measured are rare, but the odds ratio can exaggerate risk when the disease or the outcome is common (>10%).

Number Needed to Treat (NNT)- The number needed to treat is a relatively new statistical concept that attempts to relay the “real world” meaning of absolute risk reduction. It represents the number of patients who would need to be treated in order to achieve benefit in one patient. NNT is calculated as 1/ ARR. In the example above, 33 patients would need to be treated in order to prevent one death (1/ 0.03).
Paxil® (paroxetine) - On December 8, 2005, the FDA warned health care professionals that exposure to paroxetine in the first trimester of pregnancy may increase the risk for congenital malformations, particularly cardiac malformations. This warning was based on two unpublished studies, a Swedish national registry study and a U.S. insurance claims database study. These studies found that infants exposed to paroxetine in the first trimester had a 1.5- 2% risk of cardiac defect compared to 1% risk in infants in the general population or infants exposed to other antidepressants. Most of the cardiac defects observed were atrial or ventricular septal defects. The study only examined the risk of exposure in the first trimester and there is currently no data addressing risk in the later stages of pregnancy. The FDA recommends that women taking paroxetine should consult with their doctor and states that paroxetine should generally not be initiated in women who are in their first trimester of pregnancy or in women who plan to become pregnant in the near future.

Aranesp®, Epogen®, Procrit® - On December 1, 2005, the FDA notified health care professionals that these package inserts have been updated to include safety information on reports of pure red cell aplasia and severe anemia (with or without other cytopenias) associated with neutralizing antibodies to erythropoietin in patients treated with these products. Most of these reports have occurred in patients being treated for chronic renal failure.

NovoSeven® (recombinant coagulation factor VIIa) - On December 1, 2005, the FDA notified health care professionals that the safety information for this product has been updated to include reports of thrombotic and thromboembolic adverse events in non-hemophilia patients. This warning is based on the results of a study in elderly, non-hemophiliac, intracerebral hemorrhage patients which demonstrated potential increased risk of arterial thromboembolic adverse events with the use of NovoSeven, including myocardial ischemia, myocardial infarction, cerebral ischemia and/or infarction.

Long Acting Beta2- Adrenergic Agents (Advair®, Foradil®, Serevent®) – On November 18, 2005, the FDA notified the manufacturers of these products to alert health care professionals and patients that these medications may increase the risk of severe asthma episodes, and death when these episodes occur. The FDA states that these medications should not be the first line treatment for asthma and should not be used unless other medicines, including medium or low-dose corticosteroids, are unable to control the patient’s asthma. A medication guide describing this increased risk is required to be provided to patients when any of these medications are dispensed.

Ortho Evra® - On November 14, 2005, the FDA notified health care professionals and patients of new warnings that the use of Ortho Evra birth control patches exposes women to higher estrogen levels than oral birth control pills containing 35 micrograms of estrogen (about 60 percent more estrogen). Estrogen use is linked to blood clots in the legs and lungs and other clotting problems such as strokes and heart attacks. It is not known whether the higher estrogen exposure from Ortho Evra actually increases the risk of these events.

Amevive® (alefacept) - On November 10, 2005, the FDA notified health care professionals to revisions in the contraindications for Amevive, used for moderate to severe chronic plaque psoriasis. Amevive should not be administered to patients infected with HIV. Amevive reduces CD4+ T lymphocyte counts, which may accelerate disease progression or increase complications of disease in these patients.