Module 4: Key Points
Hierarchy of Evidence: Randomized Controlled Trials

- The RCT is considered one of the most powerful, if not <u>the</u> most powerful, study design for addressing therapeutic questions. RCTs are also typically characterized by other methodological features, but these are not (by design) unique to the RCT. These include:
    - Well-defined clinical question/hypothesis
    - Explicit definition of the study population
    - Use of a control group
    - Sufficient sample size to adequately address the clinical question at hand
    - Minimal loss to follow-up
    - Appropriate analytic techniques

- Since an RCT involves randomly assigning subjects to one or another treatment arms, an RCT is only justified if there exists clinical equipoise regarding the efficacy of a particular therapy or equipoise between the selection of two alternate therapies. Equipoise means that there is collective uncertainly in the medical/scientific community regarding the efficacy of a particular treatment. Equipoise has relevance not only to the selection of the active treatment arm (i.e. would not be appropriate to randomly assign subjects to receive or to not receive a therapy with proven efficacy). Equipoise is also relevant to selection of the control (placebo) group (i.e. would not be appropriate to randomly assign some subjects to receive placebo if there is a therapy with proven efficacy)

- The meaning of randomization is that treatment assignment is left to the play of chance. The goal of randomization is to control for both known and unknown confounding factors. A confounding factor is a third (or extraneous) factor (variable) that distorts the true relationship between an exposure and an outcome. Randomization ensures that such factors are evenly distributed between the treatment groups. If they are evenly distributed, then they no longer exert a confounding effect.

- Several groups of people who should be blinded – the research subject, the investigator assigning treatment and the investigator who evaluates outcome. Blinding is important (necessary) because of the potential for lack of blinding to introduce bias (systematic error). Such bias may be introduced either at a conscious or a subconscious level.

- An RCT should be designed to address/answer a very specific clinical question.
    - A primary outcome measure should be specified (e.g. incident stroke within 12 months of treatment assignment).
    - Be aware of the problem of multiple hypothesis testing (e.g. no primary hypothesis pre-specified, with the results of a multitude of post-hoc analyses presented instead)
    - Is the primary outcome measure a clinically relevant one? An RCT should address/answer a clinically relevant question.
    - Is there a discussion of the clinical significance of the difference in outcome measure that is detected between the two treatment groups? The important point

here is to differentiate clinical from statistical significance. Statistical significance can be demonstrated for even very small differences in outcome between treatment groups if the sample size is sufficiently large.

- In designing an RCT there is always a tension between the need to optimize the interval and external validity of the study. The ability to demonstrate a difference in outcome between two treatment groups requires that there is more variation between treatment groups than there is within treatment groups. Reducing within-treatment variability requires the recruitment and enrollment of a homogeneous study population. This is typically accomplished by restricting the diversity of subjects eligible to participate in the study (e.g. only patients < 60 years of age, or subjects without relevant co-morbidities such as diabetes). The problem with restricting the inclusion criteria in this way, is that the external validity (or the ability to generalize/particularize) the results of the RCT to an individual patient, is potentially compromised.

- The results of an RCT cannot be interpreted without asking whether the sample size was adequate. If the sample size was too small, then the failure to demonstrate a clinically meaningful effect of a particular treatment, may represent a type II (false negative) error. If the sample size is too small, then the demonstration of a clinically meaningful effect of a particular treatment, may represent a type I (false positive) error. One way to evaluate this latter problem is to look at the confidence interval that surrounds the point estimate of the treatment effect. If the confidence interval is broad (and especially if it spans unity), then less weight should be placed on the conclusions than if the confidence interval were narrow (and did not span unity)

- CONSORT – <u>CON</u>solidated <u>S</u>tandards <u>O</u>f <u>R</u>eporting <u>T</u>rials. CONSRT includes a flow diagram, the goal of which is to make explicit the flow of study participants (how many were recruited, randomized, lost to follow-up, included in analyses, etc.)

Limitations of a RCT
- Costly, cumbersome and time-consuming
- Restricted study population may not be generalizable to individual patients in "real life"
- Outcomes reported may not be clinically relevant
- Study procedures may not mimic "real life" and so results may not be generalizable (e.g. the benefits of CEA for carotid stenosis are critically dependent on the experience of the surgeon and low procedure-related morbidity and mortality)
- Still potential for bias (e.g. selective reporting of results)
- False negative results from underpowered studies may be misleading