Narav i obilježja kliničkih istraživanja

Trkulja, Vladimir

2006 MICC
Medical Information Conference Croatia

ministarstvo znanosti obrazovanja i športa

Ovid
a Wolters Kluwer business

CARNet
HRVATSKA AKADEMSKA I ISTRAŽIVAČKA MREŽA

The nature and characteristics of clinical research

CLINICAL RESEARCH – A quite specific field of human activity. It is a highly **administrative** and yet a highly **creative** work. When done “properly”, it is a complex mixture of scientific thinking, good ideas, reasonable hypotheses, appropriate methodology, systematic and consistent actions, enthusiastic perseverance and high ethical standards – and all this with only one goal: to generate valid and reliable data to improve human medicine

The scandal of poor medical research

“When I tell friends outside medicine that many papers published in medical journals are misleading because of methodological weaknesses they are rightly shocked....Why are errors so common? Put simply, much poor research arises because researchers feel compelled for career reasons to carry out research that they are ill equipped to perform, and nobody stops them.”

DG Altman, BMJ 1994;308:283

Meaningful clinical research

- Assures that the rights, safety and well-being of all subjects included are protected, consistent with the principles of the Declaration of Helsinki
- Assures that all collected data is credible
- Assures an appropriate data analysis and interpretation “within the true meaning of the data”, taking into account potential bias and/or limitations of the study

General practical principles

- Medical expertise is needed for clinical research – but it is not sufficient *per se*
- Scientific methodology is a “trade” as any other – it should be studied and practiced
- Not all research can or should be “revolutionary” – still there should be a certain level of “scientific relevance” – at least as a “research already done, but worth repeating”...
- In science, the “form” is equally relevant as the “content” – a great idea might be “ruined” by inappropriate “form of the research”, while a “less ingenious idea” may be a fair contribution to the overall knowledge when “materialized” in an adequate way

...and more...

- Not even a “huge” and well-designed study should be expected to answer more than one or two questions
- There is no statistical “magic” that could amend the flaws of an inappropriate design or study conductance – “garbage in – garbage out”
- A well concieved and conducted study could be ruined by inappropriate statistical analysis – “quality goods in – garbage out”

Therefore...

- **Take your time to**:
- Identify and understand a “problem” worth solving
- “Transfrom” it into a right question or two that you want to answer – define the primary objective(s)
- Consider potential APPROPRIATE ways that could lead you to your answers
- Consider FEASIBILITY - “art of compromising” is the “art of living” — review what is needed and review what is available to you – adjusting the goals in the planning phase is normal – adjusting them *post-hoc* compromises your work
- Consider the ethical aspects

...and more...

- **Construct a precise road-map that will take you to your answers**:
- Write down a precise protocol
- Consider possible obstacles and pre-define the ways to handle them
- Pay special attention to data recording
- Have in mind: every plan (protocol) is amendable – but data resulting from unpredictable “free interpretation” of the protocol is NOT
- Pre-definition of reasarch objectives, outcomes to be measured, measurement and analytical methodology is a BASIS of a meaningful research – data should shape your conclusions – and not *vice-versa*

...and more...

- **From the very beginning have in mind statistical aspects**:
- Consider the statistical properties of the study design
- Consider the sample size and the sampling procedure
- Consider the statistical properties of the data to be collected
- Consider an appropriate statistical analysis – make a statistical plan to be a part of the protocol: decide on summary statistics, on inferential statistical methods, primary and secondary analyses, possbile sensitivity analysis etc.

...finally...

- Once you have completed all this – you are ready to "start working" !

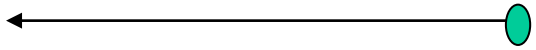
Clinical research “by nature”...

1. **Evaluates occurrence or presence of a disease (or an attitude, behavior etc.) (incidence / prevalence studies)**
2. **Evaluates etiological factors of a disease (or attitudes, behaviors etc.) or factors influencing outcomes (prognostic studies)**
3. **Evaluates a diagnostic test**
4. **Evaluates a therapeutic intervention**

Reseach by “data collection”

- **Prospective** - longitudinal or cross-sectional

- **Retrospective**



- **“Special case” – systematic review / meta-analysis**
- Raw data already collected by others “in the past” – you extract and collect them prospectively

Research by “general design”

- **Observational (“non-experimental”)** – no “intervention” is a part of the study – you just observe and collect data: **prospectively** or **retrospectivelly**

- **Experimental** – an “intervention” (therapeutic, diagnostic) is a planned part of the study – by definition – prospective.

“Designs”-by-“nature”

- **Incidence / prevalence studies** – **observational**, prospective (eg, cohort follow-up; cross-sectional) or retrospective (eg, registry data analysis etc.)
- **Prognostic studies** – **observational**, prospective (+ case-control studies) or retrospective , but account for “covariates or predictors” (may be stratified)
- **Evaluation of diagnostic tests** – **experimental**; Does a test discriminate between “disease” and “health” ? (sensitivity, specificity, ROC analysis, positive & negative predictive value); How do two tests agree ? (agreement analysis)?

...and more...

- **Evaluation of therapeutic interventions** – **experimental**
- CAUTION: single-arm treatment studies ARE NOT EXPERIMENTAL DESIGNS – requires a control group (active, placebo or at least “no-treatment”)
- Treatment comparisons
  - between subjects (parallel-group, cluster-randomized design)
  - within subjects (1st or higher-order crossover designs)
  - combined (eg, titration designs, enrichment designs, placebo-challenging designs)

...more about “therapeutic interventions”

- By “acquisition of experimental units”
  - All “at the same time” (cohort designs)
  - Group-sequential designs
- By “nature” of the outcomes
  - Pharmacokinetic (eg, relative BA or BE studies)
  - Pharmacodynamic
  - Therapeutic efficacy (and safety) or, less frequently, safety (and efficacy) studies

...and more...

- By “purpose”
- Dose-ranging studies
- “Early” efficacy-safety or “proof of the concept” studies
- Confirmatory efficacy/safety studies etc.
- By “statistical logic”
- Two-sided hypothesis study. Is there a difference between treatment A and treatment B (A may be better or worse)?; formal statistical hypothesis: H0= no difference; H1= difference
- One-sided hypothesis or superiority study. Is treatment A better (superior) than treatment B (it may be better or not)?; formal statistical hypothesis: H0= A is not better; H1= A is better

...more about “statistical logic”...

- Equivalence study. Treatment A is equivalent – neither worse, nor better – to (than) treatment B; formal statistical hypothesis: H0= inequivalence; H1= equivalence
- Non-inferiority study. Treatment A is (at least) not worse (it might be equivalent or better) than treatment B; formal statistical hypothesis: H0= inferiority; H1= non-inferiority

A few “tips” on therapeutic interventions

- Want to see whether A and B differ? – Use a “two-sided hypothesis” approach. One-sided hypothesis is rarely justified!
- Multiple comparisons, eg, A vs. B, A vs. C, B vs. C; or A vs. B on several occasions over time; or A vs. B on several (especially if related) outcomes, may lead to a “spurious” conclusion of “difference” (H0 rejection) where there is none (Type I error) – restrict the number of comparisons or adjust alpha-level (account for multiple comparisons)

...and more...

- A common mistake: Accepting H0 in a “two-sided hypothesis” test (no “statistically significant difference”) is wrongfully interpreted as no “relevant difference” or “equivalence” of A and B ! A difference (that is there) might be simply “missed” due to inadequate power (Type II error) !
- Equivalence can only be concluded in a formal equivalence test! Lack of power in an equivalence study will result in a conclusion of “equivalence not proven” – this is an “inconclusive” outcome – but at least a conclusion is not misleading...

...and more...

- The equivalence concept is typically applied when comparing interventions with the same active principle (eg, generic vs. original drug product, IR vs. SR formulation etc.)

- The non-inferiority concept is a “standard” to prove that a therapeutic intervention is “at least not worse” than a “standard” with a different active principle (eg, two beta-blockers)

Finally...

- Do not be “obsessed” by “p-values” (probability that a difference is “there by chance”) – never forget to determine size of the difference and report confidence intervals !
- A small study may fail to yield “significance”, but the size and the direction of the difference might be very informative!

In each case...

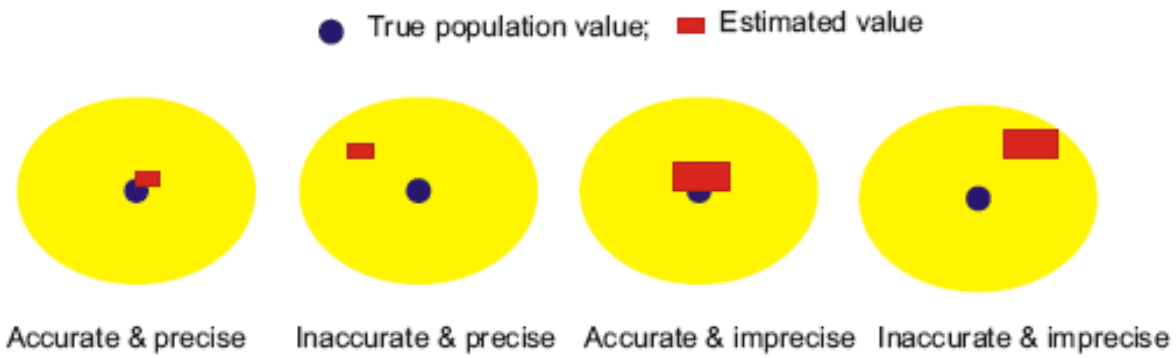
- You work with a sample (of patients, healthy subjects, urine samples, DNA samples, cultured cells...): you measure and determine *parameters*: mean values, SD, proportions, regression coefficients etc. – OR – you determine differences in parameters (between samples or factors)
- The sample, however, is not the “issue” – you do this to be able to conclude on the population
- By determining parameters in the sample – you **estimate the population parameter values** OR **population parameter differences**: How high are 10-year old boys in Zagreb? Who is higher – 10-year old boys or 10-year old girls ?

Meaningful research...

Results in **accurate** and **precise** population parameter estimates (or estimated parameter differences)

Accuracy: the estimate realiably represents the population value

Precision: the estimate is “placed” in a reasonably “narrow” range



- A study resulting in **accurate** and **precise** parameter estimates is said to have **INTERNAL VALIDITY**

Meaningful research...

- Allows you to make “broader” generalizations based on your results – “external validity”

Bias

- Leads to inaccurate and/or imprecise estimates and compromises “internal validity”
- The main types of bias:

- **Selection bias** – biased allocation or recruitment
- **Performance bias** – unequal provision of care
- **Detection bias** – outcomes not appropriate for the objective, biased measurements, inappropriate statistical analysis
- **Attrition bias** – biased occurrence and handling of protocol deviations and loss to follow-up

Poor external validity

- Sample is not representative for the population – size is not primary (inclusion/exclusion criteria not in agreement with the population properties)
- Setting is not “representative” for the “real life” – e.g., standards of care “suboptimal” etc., measurement methods differ from the “real life”..

Ascertain validity of your research

GENERAL PRINCIPLES – CAREFULL PLANNING ESSENTIAL

- Choose a right design – prospective studies are more likely to provide credible data (more complete, reliably determined...)
- Criteria for “enrollment” of “units” – the sample should be representative for the population
- Selection of outcomes – do not exaggerate in number; measured outcomes must represent the property you want to evaluate
- Sample size – use own or “other” experience
- Protocol procedures – should “comply” with the “real life”
- “Stick” to the protocol

...by type of study...

- **Incidence (or prevalence studies)**; eg, *Dobec-Meić, Pikija, Cvetko, Trkulja et al. J Neurooncol 2006*
- Objective: determine incidence of primary intracranial tumors in the Varaždin County 1994-2004 by type; compare men and women
- Design: retrospective population-based incidence study
- INTERNAL VALIDITY – good:

- Population base= 2001 Croatian census data + precise criteria for “cases” (selection bias)
- Do not miss cases ! Searched all hospitals in Zagreb and Varaždin (attrition bias)
- Compare men and women as IRR with Poisson-based CI (detection bias)

EXTERNAL VALIDITY – moderate – informative for the County, not informative for Croatia

...and more...

- **Prognostic studies** – eg, *Pikija, Milevčić, Trkulja et al. Eur Neurol 2006*
- Objective: Does fasting serum Tg at admission predict ischemic CVI size ?
- Design: prospective prognostic study
- INTERNAL VALIDITY - good:

- Consecutive first-ever ischemic CVI patients admitted within 24 hrs since symptom onset over a pre-defined period of time (1 year) (selection bias);
- All undergo the same “in-house” algorithm for ishemic CVI (performance bias);
- CVI volume determined by a validated method based on CT-scans by a blinded investigator; multiple regression analysis (detection bias)
- All assessed for other relevant factors: cholesterol, glucose, blood pressure, atrial fibrillation, type of CVI, timing of CT-scan relative to admission (attrition/performance bias)

EXTERNAL VALIDITY - good – “in house” algorithm and measurement methods in line with international “standard of care”, in line with SSS

...and more...

- **Evaluation of a diagnostic test** – eg, *Mađarević, Pećina, Trkulja et al.* in preparation
- Objective: Does 2nd metatarsal cortical index “recognize” forefoot overload ?
- Design: prospective
- INTERNAL VALIDITY – good:

- “cases” defined by clinical metatarsalgia and pedobarographic measurement, “healthy” – age, sex and BMI-matched controls (selection bias)
- Forefoot X-rays A-P and L-L, analyzed by a validated software (performance bias)
- X-ray analysis “blinded”, data analysis “blinded”, optimal cut-off value by ROC analysis, positive and negative predictive values determined (detection bias)

EXTERNAL VALIDITY – moderate – restricted to female patients 20-40 years of age (most commonly affected) – may not be “extrapolated” to men, or to other age-groups

Finally...

- **Therapeutic interventions** – eg, *Milutinović, Plavljanić, Trkulja. Fund Clin Pharmacol 2006*
- Objective: Is there an efficacy difference between two epoetin brands in treatment of renal anemia ?
- Design: Two single-center prospective parallel-group trials
- INTERNAL VALIDITY – good:

- Randomization and treatment-allocation concealment (selection bias)
- Single-blinded, blinded data review (ITT and PP data sets) (selection, performance, detection, attrition bias)
- Precise protocol procdures, no protocol violations (peformance bias)
- Primary outcomes Hb values and epoetin doses, analysis of covariance ITT and PP, accounting for baseline Hb, iron parameters, CRP, dialysis dose, age, sex, ACE inhibitors, iPTH, HD duration (detection bias, attrition bias)

EXTERNAL VALIDITY – moderate – outcomes valid; measurements standard, validated, inclusion/exclusion critaria typical for “epoetin trials” – but single-center

To conclude...

- Clinical research is logistically demanding...but more than that, it demands methodological knowledge and common sense – one without the other “does not work”
- Living in a small, financially underpriviledged and scientifically “annonymous” country like Crotia does not mean that meaningful clinical research is impossible...
- However...
- “We need less research, better research, and research done for the right reasons.....”

DG Altman, BMJ 1994;308:283