CLINICAL TRIAL DESIGN
AND
CLINICAL EVIDENCE

Prof. dr. A. Dupont,
Dept. of Clinical Pharmacology and Therapeutics
RATIONAL PHARMACOTHERAPY

- Nairobi Declaration: “The right drug – in the right dose – for the right patient – at affordable cost”

- Use only medicines of which the efficacy and effectiveness have been proven in validated clinical trials

- Use the medicines in a correct way: indication(s), contra-indication(s), dose(s), interactions, side effects, variability (genetic polymorphism...)

- Pay attention to the cost of the treatment

- Need for EVIDENCE
EVIDENCE-BASED MEDICINE

“The conscientious, explicit and judicious use of current best evidence in making decisions about the care of individual patients”

“The practice of evidence-based medicine means integrating individual clinical expertise with the best available external clinical evidence from systematic research”
EVIDENCE-BASED PHARMACOTHERAPY

- The strength of the evidence
- The quality of the studies
- The interpretation
  - surrogate versus clinical endpoints
  - absolute versus relative changes
  - number needed to treat
Objective: assessment of benefit/risk ratio

→ CONTROLLED CLINICAL TRIAL:

to allow for a valid comparison between different types of treatment in otherwise similar groups
OBSERVATIONAL STUDIES

Retrospective or prospective

✓ *cohort-studies*: cause to consequence studies

✓ “*case-control*” (“patient-control”) studies

P.S. “Bias, confounding”
RANDOMIZED TRIALS

Always prospective

✓ control group
✓ parallel or “cross–over”
✓ open or blind (“single, double, triple”)
✓ randomization

RANDOMIZED DOUBLE-BLIND CONTROLLED TRIAL

= golden standard
BIAS

= all factors of design, conduct, analysis and interpretation of the results of clinical trials that lead to an estimate of a treatment effect which deviates from its true value

[= anything which could affect patient response other than the difference between the treatments]
CLINICAL TRIALS: methods to minimize bias

- **Randomization**
  - to assure comparability of test groups
  - to minimize selection bias

- **Blinding**
  - double-blind to assure that subjective assessments are not affected by knowledge of treatment assignment

- **Compliance**
  - document actual usage of the drug
Observational studies: methods to minimize bias

• Observational studies can be designed with rigorous methods that mimic those of clinical trials

• "Restricted cohort" design (Horwitz et al, 1979)
  
  - identifies “zero time” for determining a patient’s eligibility
  - uses inclusion and exclusion criteria similar to those of clinical trials
  - adjusts for differences in baseline susceptibility to the outcome
  - use statistical methods (e.g., intention-to-treat analysis) similar to those of clinical trials
CONTROL GROUPS

- concurrent placebo control
- “no treatment” concurrent control
- dose – response control
- active (positive) concurrent control
- external control (incl. historical control)
- combination of control groups (e.g. *three-arm* study)
Purpose of a control group

→ to allow discrimination of outcomes in patients caused by the test drug from outcomes caused by other factors (e.g. natural progression of the disease, observer expectations, other treatment…)

✓ in most situations it is not possible to predict outcome from the baseline and treatment variables or to identify a population that is similar to the test population with respect to all relevant variables

→ **concurrent control group is needed**
  (chosen from the same population as the test group)
CHOICE OF CONTROL GROUP IN CLINICAL TRIALS

affects

✓ the inferences that can be drawn from the trial
✓ the degree to which bias can be minimized
✓ the kind of subjects that can be recruited
✓ the public credibility of the results
✓ the kind of endpoints that can be studied
✓ the time needed to carry out the trials
✓ the conduct and the interpretation of the study
Purpose of clinical trials
→ to show efficacy and/or safety

✓ by demonstrating “superiority” over the control (placebo, lower dose, active drug)
✓ by demonstrating that the new drug is similar, or not inferior in efficacy to an established effective dose of another therapy (active control) “equivalence” trials, or – more strictly – “non-inferiority” trials

To show the relative (comparative) efficacy, safety, benefit/risk relationship or utility of two treatments (problems: choice of control, choice of dose)
Trials to show “equivalence” or “non-inferiority”

• Equivalence trial :
  – finding of interest is two-sided,
    e.g. bioequivalence trials
    clinical equivalence trials
    (generic versus original product, which is not absorbed)

• Non-inferiority trials :
  – to show that the efficacy of investigated product
    is no worse then that of the active comparator;
    finding of interest is one-sided

(“is the new drug within a defined margin of the active control?”)
Assay sensitivity / Trial sensitivity

• Superiority trial
  The finding of a difference (superiority) itself documents assay/trial sensitivity (validity) and documents the efficacy of the superior treatment

• Non-inferiority trial
  If finding of non-inferiority: two remaining questions
  (1) was the study capable of detecting a difference if there really is one? (did the trial have “assay sensitivity”?)
  (2) do historical data indicate that the active control is better than placebo?
    (no “internal-standard”, except with 3-arm study with placebo)
Difference

Significant

true: difference is real

useful
interest
clinically not relevant

false
sampling error
(type 1 error,
acceptable probability = $\alpha$)
bias
(distorted comparison)

Non-significant

treatments equivalent

difference not detected

sampling error
(type 2 error,
acceptable probability = $\beta$)
comparison not sensitive

too few subjects
unfavorable conditions
No statistically significant difference = equivalence ??

→ NOT TRUE!

→ Not reaching statistical significance in a clinical trial is often due to there being insufficient subjects in the study

★ Concluding equivalence or non-inferiority based on observing a non-significant test result of the null hypothesis that there is no difference between the investigational product and the active comparator is inappropriate.
Trial execution quality

• Superiority trial
  - there is a strong imperative to take steps to reduce the likelihood of failing to show a difference between treatments when one exists

• Trials intended to show no-difference between two treatments
  - there is much weaker incentive to engage in attempts to reduce error and assure assay sensitivity
TYPES OF EVIDENCE

- **“Efficacy”**: clinical trials
- **“Effectiveness”**: drug use in daily practice
- **“Efficiency”**: cost-effectiveness
TYPES OF EVIDENCE

- **Surrogate endpoints versus clinical endpoints**
  - quality of life, morbidity, hospitalization, mortality

- **Short term versus long term effects**

- **Primary versus secondary endpoints**

- **Quantitative evaluation of the effects**
  - absolute versus relative risk reduction
  - “Number Needed to Treat (NNT)”
  - “Number Needed to Harm (NNH)”
PRIMARY and SECONDARY ENDPOINTS

Primary endpoint:

- should be able to provide relevant clinical information
- should reflect accepted norms in the therapeutic area
- is used for estimating sample size
- usually relates to efficacy but can on occasions be related to safety, quality of life, health economics

ICH:
‘There should generally be only one primary endpoint’
Secondary variables

- either supportive measurements related to the primary objective or measurements of effects related to the secondary objectives

- pre-definition in the protocol is also important, as well as an explanation of their relative importance and roles in interpretation of trial results

- number of secondary variables should be limited and should be related to the limited number of questions to be answered in the trial
CLINICAL versus SURROGATE ENDPOINTS

• Optimal primary endpoint is a clinical event or outcome (e.g. death, myocardial infarction, stroke, prolonged survival of a patient with heart failure, …)

• Demonstrating a statistically significant effect on outcomes requires large and expensive trials of long duration (e.g. mortality trials in hypercholesterolemia, hypertension, CHF, stroke, …)

• Need for evidence from early exploratory trials and confidence that further development of the pharmaceutical product is warranted
A **surrogate** endpoint (marker) has been defined by the FDA (Dr. Robert Temple) as:

- “A laboratory measurement or a physical sign used as a substitute for a clinically meaningful endpoint that measures directly how a patient feels, functions, or survives.

- *Changes induced by a therapy on a surrogate endpoint are expected to reflect changes in a clinically meaningful endpoint.*
<table>
<thead>
<tr>
<th>Disease</th>
<th>Surrogate marker</th>
<th>Clinical Endpoint</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cardiovascular</td>
<td>Blood Pressure</td>
<td>Myocardial infarction</td>
</tr>
<tr>
<td>Cardiovascular</td>
<td>Cholesterol</td>
<td>Coronary Heart disease</td>
</tr>
<tr>
<td>Cancer (solid tumor)</td>
<td>Tumor response</td>
<td>Mortality</td>
</tr>
<tr>
<td>Osteoporosis</td>
<td>Bone Mineral density</td>
<td>Fractures</td>
</tr>
<tr>
<td>Diabetes</td>
<td>Blood glucose</td>
<td>Diabetes complications</td>
</tr>
</tbody>
</table>
NON-VALIDATED SURROGATE ENDPOINTS can lead to mistakes

✓ post-infarction patients with a high rate of ventricular extrasystoles (VES) have a higher mortality risk
✓ the higher incidence in sudden cardiac death was attributed to VES
✓ it was assumed that reduction of VES with anti-arrhythmics would reduce post myocardial infarction mortality

✓ but: Cardiac Arrhythmia Suppression Trials (CAST I and II) showed 24% increase in death rate in the active treated group as compared to placebo, invalidating VES as a reliable predictor and surrogate marker
RELATIVE versus ABSOLUTE EFFECT
adjusting for baseline risk

e.g. SHEP (Systolic Hypertension in Elderly) trial
“CHD-events” (incidence over 5 years)
active treated group 4.4 %
placebo group 5.8 %

→ relative risk 0.73
relative risk reduction 27 %
absolute risk reduction 1.4 %
14 “events” avoided for every 1000 patients
1 “event” avoided for every 71 patients treated (“NNT”)
“NUMBER NEEDED TO TREAT (NNT) – NUMBER NEEDED TO HARM (NNH)”

NNT = How many patients must be treated to obtain the desired outcome in one?
NNT = 1 / ARR or 1 / (ARC – ART)

Suppose a medicine that reduces 1 year mortality by 40% 

- high background risk (yearly mortality 1/10)
  ARR: 0.04 (0.4 x 0.1); NNT 25 (1/0.04)

- low background risk (yearly mortality 1/100)
  ARR: 0.004 (0.04 x 0.1); NNT 250 (1/0.004)

NNH = How many patients can be treated before an adverse effect of the treatment occurs?
## ABSOLUTE versus RELATIVE RISK REDUCTION

<table>
<thead>
<tr>
<th></th>
<th>RRR %</th>
<th>ARR %</th>
<th>“NNT”</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>WOSCOPS</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>low risk</td>
<td>– 30</td>
<td>– 1,5</td>
<td>66</td>
</tr>
<tr>
<td>moderate risk</td>
<td>– 30</td>
<td>– 3,0</td>
<td>33</td>
</tr>
<tr>
<td>higher risk</td>
<td>– 30</td>
<td>– 4,5</td>
<td>22</td>
</tr>
<tr>
<td>high risk</td>
<td>– 30</td>
<td>– 6,0</td>
<td>17</td>
</tr>
<tr>
<td>total</td>
<td>– 30</td>
<td>– 2,5</td>
<td>40</td>
</tr>
<tr>
<td><strong>4S–Study</strong></td>
<td>– 34</td>
<td>– 7,9</td>
<td>13</td>
</tr>
<tr>
<td>LEVEL</td>
<td>TYPE</td>
<td></td>
<td></td>
</tr>
<tr>
<td>-------</td>
<td>------</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Ia</td>
<td>Evidence based on a systematic review or meta-analysis of randomized clinical trials (RCT)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Ib</td>
<td>Evidence based on at least 1 RCT</td>
<td></td>
<td></td>
</tr>
<tr>
<td>IIa</td>
<td>Evidence based on at least 1 well designed controlled study without randomization</td>
<td></td>
<td></td>
</tr>
<tr>
<td>IIb</td>
<td>Evidence based on at least 1 other well designed experimental study</td>
<td></td>
<td></td>
</tr>
<tr>
<td>III</td>
<td>Evidence based on at least 1 well designed non-experimental study (comparative study, correlation study, case study)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>IV</td>
<td>Evidence based on report of an expert and/or opinion and/or clinical experience of “respected authorities”</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Gradation</td>
<td>Evidence</td>
<td></td>
<td></td>
</tr>
<tr>
<td>-----------</td>
<td>----------</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>A</strong></td>
<td>Minimum 1 RCT as part of a well performed study of literature (evidence level Ia &amp; Ib)</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>B</strong></td>
<td>Well designed clinical trials, but no RCTs (evidence level IIa, IIb &amp; III)</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>C</strong></td>
<td>Expert committee report and/or opinion and/or clinical experience of respected authorities in the absence of good studies (evidence level IV)</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Randomized controlled trials, observational studies and the hierarchy of research designs


- “average results from well–designed observational studies (with a cohort or case–control design) do not systematically overestimate the magnitude of the associations between exposure and outcome as compared with results of randomized controlled trials”

- “observational studies may be less prone to heterogeneity in results” (include a broader representation of the population than RCTs)
INDIVIDUAL CLINICAL TRIALS

1. Objectives: should be clearly stated
2. Design
   ✓ e.g. parallel group/cross-over
   factorial
dose escalation
fixed dose–response
   ✓ appropriate comparators
   ✓ appropriate number of subjects
   ✓ primary/secondary endpoints
   ✓ analysis plan
   ✓ methods of monitoring adverse events
   ✓ procedures for follow-up of drop outs
3. Conduct
4. Analysis  description of statistical methods
5. Reporting