Study design continued: intervention studies

Gustaf Edgren, PhD
Karolinska Institutet

Outline

- Background
- Uncontrolled trials
- Non-randomized, controlled trials
- Randomized controlled trials
  (with an overview on how to assess an RCT paper)

Repetition – principal approaches

Experiments	**Observational studies**
”Change the state of nature and observe the effects”	”Observe nature as it is”
Examples: - Randomized studies - Cross-over studies - Etc.	Examples: - Cohort studies - Case-control studies - Cross sectional studies

Background

- We mentioned earlier today that uncontrollable confounding is one of the principal limitations of observational studies
  - Often there is a relation between why an individual chooses a certain behavior/treatment/etc and their health status
- We therefore need alternatives to controlling this confounding

Notes on intervention studies

- All intervention studies are extensions of the prospective cohort framework
- The key difference being that exposure is actively allocated by the investigator
- The analysis and interpretation otherwise follows the same general principles

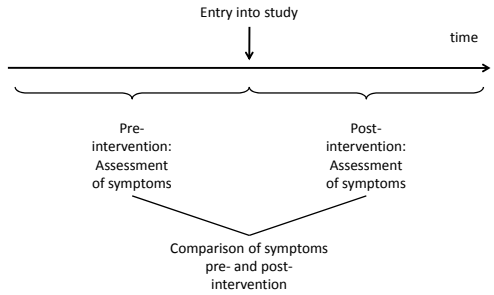
Types of intervention trials

- Uncontrolled trials
- Controlled (non-randomized) trials
- Randomized controlled (clinical) trials
- Variations…

Uncontrolled trials

- In an uncontrolled trial, the symptom levels or risk of the outcome after the initiation of a certain treatment is compared to historic levels
- Each study participant is thus compared to themselves, before initiation of treatment
- Rarely a suitable design, but can sometimes be an acceptable compromise for incurable or very predictable conditions (ALS, pancreatic cancer, CHF NYHA IV, etc.)

Uncontrolled trials



Uncontrolled trials – example

Inhaled Iloprost To Treat Severe Pulmonary Hypertension

An Uncontrolled Trial

Horst Olschewski, MD; H. Ardeschir Ghofrani, MD; Thomas Schmehl, PhD; Jörg Winkler, MD; Heinrike Wilkens, MD; Marius M. Höper, MD; Jürgen Behr, MD; Franz-Xaver Kleber, MD; and Werner Seeger, MD, for the German PPH Study Group*

Ann Intern Med. 2000;132:435-443.

Uncontrolled trials – example

Results: During the first 3 months of therapy, New York Heart Association functional class improved in 8 patients and was unchanged in 7 patients. Four patients died, 3 of right-heart failure and 1 of sepsis. The acute hemodynamic response to inhaled iloprost was predominant pulmonary vasodilatation with little systemic effect at baseline and at 3 months (data available for 12 patients). Hemodynamic variables were improved at 3 months, and the distance walked in 6 minutes improved by 148 m (95% CI, 4.5 to 282 m; $P = 0.048$). Of the 15 patients who continued to use inhaled iloprost, 8 stopped: Four had lung transplantation, 1 switched to intravenous prostacyclin therapy, and 3 died. Seven patients are still receiving inhaled iloprost (mean ± SD) duration of therapy, 536 ± 309 days; mean dosage, 164 ± 38 μg/d).

Conclusions: Inhaled iloprost may offer a new therapeutic option for improvement of hemodynamics and physical function in patients with life-threatening pulmonary hypertension and progressive right-heart failure that is refractory to conventional therapy.

Ann Intern Med. 2000;132:435-443.

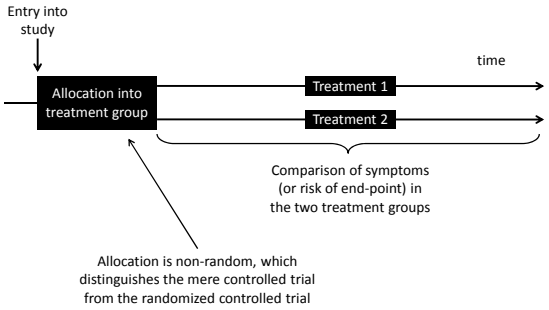
Uncontrolled trials – problems

- Natural disease variation
- Spontaneous cure/improvement
- Placebo! (patient expectations and hopes)
- Investigator hope and wishes

Controlled (non-randomized) trials

- Instead of comparing symptoms before and after initiation of treatment, the comparison is made between patient groups receiving different treatment
- The allocation to treatment groups is non-random (e.g. by patient or physician choice)
- Can be the only alternative when randomization is not possible, or even ethical (i.e. no *equipoise**)

Controlled (non-randomized) trials



Controlled trials – example

JOURNAL OF BONE AND MINERAL RESEARCH
Volume 20, Number 4, 2005
Published online on December 6, 2004; doi: 10.1359/JBMR.041203
© 2005 American Society for Bone and Mineral Research

Treatment of Painful Vertebral Fractures by Kyphoplasty in Patients With Primary Osteoporosis: A Prospective Nonrandomized Controlled Study

Christian Kasperk,[1,2] Jochen Hillmeier,[2,3] Gerd Nöldge,[4] Ingo A Grafe,[1] Katharina DaFonseca,[1,3] Dorothea Raupp,[3] Hubert Bardenheuer,[5] Martin Libicher,[4] Ute Monika Liegibel,[1] Ulrike Sommer,[1] Ulrike Hilscher,[1] Walter Pyerin,[6] Marcus Vetter,[6] Hans-Peter Meinzer,[6] Peter-Jürgen Meeder,[7] Rod S Taylor,[7] and Peter Nawroth[1]

ABSTRACT: This study investigates the effects of kyphoplasty on pain and mobility in patients with osteoporosis and painful vertebral fractures compared with conventional medical management.

Introduction: Pharmacological treatment of patients with primary osteoporosis does not prevent pain and impaired activity of patients with painful vertebral fractures. Therefore, we evaluated the clinical outcome after kyphoplasty in patients with vertebral fractures and associated chronic pain for >12 months.
Materials and Methods: Sixty patients with primary osteoporosis and painful vertebral fractures presenting for >12 months were included in this prospective, nonrandomized controlled study. Twenty-four hours before performing kyphoplasty, the patients self-determined their inclusion into the kyphoplasty or control group so that 40 patients were treated with kyphoplasty, whereas 20 served as controls. This study assessed changes in radiomorphology, pain visual analog scale (VAS) score, daily activities (European Vertebral Osteoporosis

Controlled trials – problems?

- Difficult to allocate patients to the different treatments with equal distribution of prognosis/symptoms
- The patient's or investigator's choice of allocation is often related to the patient's prognosis or symptoms
  - Healthier patients (and physicians of healthier patients) typically want more active treatment, etc.
- Consequentially, we choose to distribute the patients randomly…

Randomized controlled trials

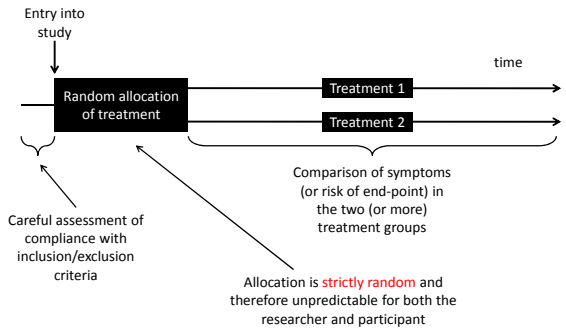
(Sometimes referred to as randomized clinical trials)

- An extension of the common controlled trial (which, as we've already concluded is an extension of the prospective cohort)
- Patients are allocated to different treatment groups in a random fashion
  - Thereby creating comparable exposure groups, with perfect (?) control of all known (and unknown) confounding factors

Randomized controlled trials

- Beyond Lind's neat experiments in the 1700's, the use of randomization in science was pioneered in the Streptomycin in TB study
- The use of sealed envelopes with random numbers was proposed by Sir Austin Bradford Hill (inspired by agricultural research done by Sir Ronald Fisher)

Randomized controlled trials



Note on placebo

- Typically, in randomized controlled trials, a "new" treatment is compared to placebo
- Often, the comparison group receives placebo (sugar pills), but if an established treatment exists, the use of placebo is (considered) unethical
- The use of placebo treatment is by no means limited to pharmacology studies, but has an important role also in for example surgical studies
- In any case, the choice of placebo must reflect best available, established treatment
- While the placebo effect can be profound, for hard outcomes, it should not be exaggerated*

*Hrobjartsson A *et al*. N Engl J Med, 2001.

Clinical equipoise

- Conversely to the choice of placebo, where the comparison group must be the best established treatment, there must be genuine uncertainty regarding the benefits of the new treatment
- A state of equipoise must persist throughout the conduct of the study
- If one treatment shows superior to the other before the study is finished, it must be terminated

Randomized controlled trials – why?

"Other study designs, including non-randomized controlled trials, can detect associations between an intervention and an outcome. But they cannot rule out the possibility that the association was caused by a third factor linked to both intervention and outcome"

Sibbald *et al*, BMJ 1998;316:201

Randomized controlled trials – good example!

The New England
Journal of Medicine

Copyright © 2002 by the Massachusetts Medical Society

VOLUME 347 JULY 11, 2002 NUMBER 2

A CONTROLLED TRIAL OF ARTHROSCOPIC SURGERY FOR OSTEOARTHRITIS OF THE KNEE

J. Bruce Moseley, M.D., Kimberly O'Malley, Ph.D., Nancy J. Petersen, Ph.D., Terri J. Menke, Ph.D., Baruch A. Brody, Ph.D., David H. Kuykendall, Ph.D., John C. Hollingsworth, Dr.P.H., Carol M. Ashton, M.D., M.P.H., and Nelda P. Wray, M.D., M.P.H.

Designing (or reading papers on) RCTs

- 10 step RCT appraisal scheme
  1. Relevant (and explicit) hypothesis
  2. Randomization process
  3. Success of randomization
  4. Fair comparison
  5. Sufficient/successful blinding
  6. Loss to follow-up/intention to treat
  7. Clinical relevance of findings
  8. Power (especially for null findings)
  9. External validity

1. Hypothesis

- It is a fair requirement that unless explicit, all studies should be hypothesis driven (and not hypothesis generating).
  - You should therefore expect to be able to find an explicit mention of what the hypothesis or aim of the study was
- More importantly, does the study design match the hypothesis
  - I.e. are the researchers answering the question they are posing?

2. Randomization process

- The goal of randomization is to make sure that all confounding factors (known or unknown) are equally distributed between comparison groups
- Furthermore, to successfully eliminate any inequalities between treatment groups, randomization must be unpredictable for patients and clinicians alike, and impossible to influence
- Why?

2. Randomization process, cont.

- Lured by promises of new drugs, patients and clinicians alike may only want to join a new study if they get allocated to the new treatment group.
- Since patients with such wishes may (or are likely to) have a different prognosis, randomization must be unpredictable
- Therefore, schemes such as "odd or even birth date," which may seem perfectly random, are ill advised.

3. Randomization success

- Techniques like sequential sealed envelopes with random numbers, or phone allocation are used in most clinical trials
- These techniques are usually fine, but were they successful in this particular trial?
  - Check Table 1 for differences in background variables between the different treatment groups
  - If there are more significant differences than expected (i.e. >5%) – be suspicious!

Successful randomization?

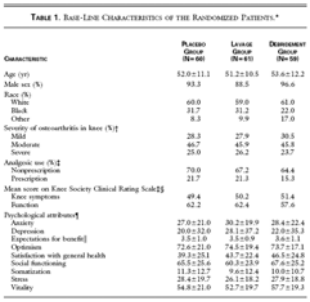


TABLE 1. BASE-LINE CHARACTERISTICS OF THE RANDOMIZED PATIENTS.*

CHARACTERISTIC	PLACEBO GROUP (N=60)	LAVAGE GROUP (N=61)	DEBRIDEMENT GROUP (N=58)
Age (yr)	52.0±11.1	51.2±10.5	53.6±12.2
Male sex (%)	93.3	88.5	96.6
Race (%)			
White	60.0	59.0	61.0
Black	31.7	31.2	22.0
Other	8.3	9.9	17.0
Severity of osteoarthritis in knee (%)†			
Mild	28.3	27.9	30.5
Moderate	46.7	45.9	45.8
Severe	25.0	26.2	23.7
Analgesic use (%)‡			
Nonprescription	70.0	67.2	64.4
Prescription	21.7	21.3	15.3
Mean score on Knee Society Clinical Rating Scale‖§			
Knee symptoms	49.4	50.2	51.4
Function	62.2	62.4	57.6
Psychological attributes¶			
Anxiety	27.0±21.0	30.2±19.9	28.4±22.4
Depression	20.0±32.0	28.1±37.2	22.0±31.3
Expectations for benefit‖	3.5±1.0	3.5±0.9	3.6±1.1
Optimism	72.6±21.0	74.5±19.4	73.7±17.1
Satisfaction with general health	39.3±25.1	43.7±22.4	46.5±24.8
Social functioning	65.5±25.4	60.3±23.9	67.6±25.2
Somatization	11.3±12.7	9.6±12.4	10.0±10.7
Stress	28.4±19.7	26.1±18.2	27.9±18.8
Vitality	54.8±21.0	52.7±19.7	57.7±19.3

4. Fair comparison

- A classic trick by pharmaceutical companies in comparison of a new drug to the best available treatment is to give the optimal dose of the new drug and a lower than normal dose for the placebo group
- Similarly, the placebo group may be composed of antiquated drugs or treatments
- This type of error may be quite difficult to spot unless you're familiar with the drugs, but can be very important

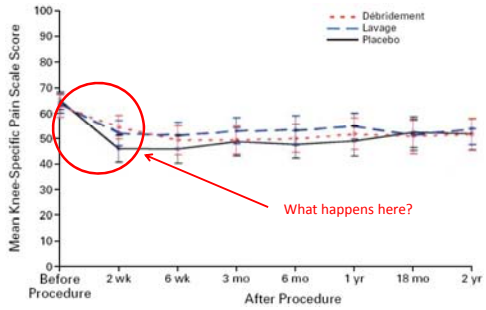
5. Success of blinding

- In addition to successful randomization, the treatment allocation should (where possible) be blind to the patients and investigators
- This process (called blinding) is used to equalize the effects of randomization
- Blinding can be used not only for medical trials, sham surgery makes this an alternative also for surgery
- Beyond the obvious (pills looking the same, etc.), blinding should also involve equal access to auxiliary treatment (such as follow-up visits and other diagnostic procedures, etc.)

5. Success of blinding, cont.

- Variants of blinding exists, and may be applied where relevant
  - Double blind – Both the patient and investigator is unaware of the treatment the patient is receiving
  - Single blind – only the patient is unaware of the treatment status
  - Triple blind – similar as a double blind trial, but the blinding is maintained also through the analysis phase

Motivation for blinding



6. Loss to follow-up/intention to treat

- In all trials, patients will leave the trial. This may be due to death or just choice.
- While the strife, naturally, should be to keep dropout minimal, it is the handling of dropout that matters
  - If dropout is significant and unequal, it is important to keep track of why patients left
- The analysis, should only consider the treatment status the patient was randomized to
  - Intention to treat

6. Intention to treat

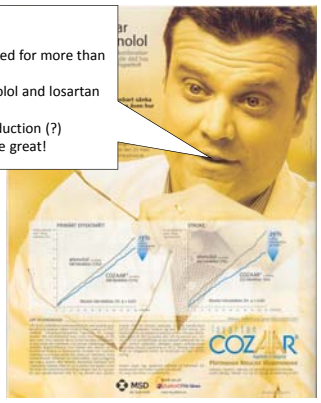
- The intention to treat is an important paradigm whereby patients are analyzed as randomized, not as treated
  - Thus, in ITT analysis, no heed is given to patient crossover between treatment groups
  - Although counterintuitive, ITT ensures minimal bias from differential dropout and crossover due to e.g. side-effects
  - Further, ITT ensures that the trial tests not the efficacy of the treatment itself, but rather the treatment policy – mimicking clinical reality

7. Clinical relevance of findings

- In some medical specialties*, researchers have to conduct enormous trials to tease apart the minute benefits of the different treatments
- Such trials may deliver very small, yet significant findings that may not be relevant in clinical practice
- A useful tool to assess this is numbers needed to treat

*Cardiology anyone?

What?
Behemoth RCT
9193 patients followed for more than five years
Comparison of atenolol and losartan
Resultat?
25 % relative risk reduction (?)
P = minute, therefore great!



8. Power

- Absence of evidence is not evidence of absence!
- In studies showing null findings, one must keep in mind whether the study was large enough to detect a meaningful difference*
- *Power (which is defined as the probability that a difference of a certain magnitude can be detected given a certain sample size) will be covered on Friday*

*Applicable to all designs

9. External validity

- The external validity or generalizability of a study concerns whether the study is applicable to other settings
- Often exclusion/inclusion criteria tend to exclude "normal" patients from studies
- Suitable questions could be:
  - Are these patients representative of "my" patient?
  - Could my patient have participated in this study?
- Keep in mind that a poor external validity does not directly disqualify a study. After all, we do generalize findings from mice…

Summary – intervention trials

- Intervention studies are natural extensions of the prospective cohort where exposure is allocated by the investigator
- Whereas the more basic intervention studies offer little advantage to cohort studies, randomization is an efficient way of eliminating confounding

Summary – randomized clinical trials

- The process of randomization must be completely random in order for all confounding factors to be equally distributed
- Further, randomization must be
  - Unpredictable, and
  - Impossible to influence
- The randomization must also be coupled with suitable analysis according to the intention to treat principle