Replicated N-of-1 RCTs for Rare Diseases

Patrick Onghena

Faculty of Psychology and Educational Sciences
KU Leuven, Belgium

Webinar 30 June 2023

KU LEUVEN

EUROPEAN JOINT PROGRAMME RARE DISEASES

SPECIAL ARTICLE

DETERMINING OPTIMAL THERAPY — RANDOMIZED TRIALS IN INDIVIDUAL PATIENTS

GORDON GUYATT, M.D., DAVID SACKETT, M.D., D. WAYNE TAYLOR, M.Sc., JOHN CHONG, M.D., ROBIN ROBERTS, M.Sc., AND STEWART PUGSLEY, M.D.

Abstract Although the treatment of an individual patient in routine clinical practice has been likened to an experiment, the method is so susceptible to bias that we have come to demand multi-patient, double-blind, randomized controlled trials on matters of efficacy. Unfortunately, such trials have not or cannot be carried out for many clinical disorders; even when they have been executed their results may be difficult to extrapolate to individual patients.

To resolve this problem, we have begun to use double-blind randomized trials in which a single patient undergoes a series of pairs of treatments, consisting of one active and one placebo or alternative treatment per pair, with the order determined by random allocation. Appropriate treatment targets (signs, symptoms, or laboratory tests) are used as the measure of efficacy, and the trial is continued until efficacy is established or disproved. We describe such a trial, which resulted in a dramatically beneficial modification of treatment in a patient with partially reversible airflow limitation. We have established a clinical service that facilitates the widespread use of the method in our community. (N Engl J Med 1986; 314:889-92.)

SPECIAL ARTICLE

DETERMINING OPTIMAL THERAPY — RANDOMIZED TRIALS IN INDIVIDUAL PATIENTS

GORDON GUYATT, M.D., DAVID SACKETT, M.D., D. WAYNE TAYLOR, M.Sc., JOHN CHONG, M.D., ROBIN ROBERTS, M.Sc., AND STEWART PUGSLEY, M.D.

Abstract Although the treatment of an individual patient in routine clinical practice has been likened to an experiment, the method is so susceptible to bias that we have come to demand multi-patient, double-blind, randomized controlled trials on matters of efficacy. Unfortunately, such trials have not or cannot be carried out for many clinical disorders; even when they have been executed their results may be difficult to extrapolate to individual patients.

To resolve this problem, we have begun to use double-blind randomized trials in which a single patient undergoes a series of pairs of treatments, consisting of one active and one placebo or alternative treatment per pair, with the order determined by random allocation. Appropriate treatment targets (signs, symptoms, or laboratory tests) are used as the measure of efficacy, and the trial is continued until efficacy is established or disproved. We describe such a trial, which resulted in a dramatically beneficial modification of treatment in a patient with partially reversible airflow limitation. We have established a clinical service that facilitates the widespread use of the method in our community. (N Engl J Med 1986; 314:889-92.)

SPECIAL ARTICLE

DETERMINING OPTIMAL THERAPY — RANDOMIZED TRIALS IN INDIVIDUAL PATIENTS

GORDON GUYATT, M.D., DAVID SACKETT, M.D., D. WAYNE TAYLOR, M.Sc., JOHN CHONG, M.D., ROBIN ROBERTS, M.Sc., AND STEWART PUGSLEY, M.D.

Abstract Although the treatment of an individual patient in routine clinical practice has been likened to an experiment, the method is so susceptible to bias that we have come to demand multi-patient, double-blind, randomized controlled trials on matters of efficacy. Unfortunately, such trials have not or cannot be carried out for many clinical disorders; even when they have been executed their results may be difficult to extrapolate to individual patients.

To resolve this problem, we have begun to use double-blind randomized trials in which a single patient undergoes a series of pairs of treatments, consisting of one active and one placebo or alternative treatment per pair, with the order determined by random allocation. Appropriate treatment targets (signs, symptoms, or laboratory tests) are used as the measure of efficacy, and the trial is continued until efficacy is established or disproved. We describe such a trial, which resulted in a dramatically beneficial modification of treatment in a patient with partially reversible airflow limitation. We have established a clinical service that facilitates the widespread use of the method in our community. (N Engl J Med 1986; 314:889-92.)

SPECIAL ARTICLE

DETERMINING OPTIMAL THERAPY — RANDOMIZED TRIALS IN INDIVIDUAL PATIENTS

GORDON GUYATT, M.D., DAVID SACKETT, M.D., D. WAYNE TAYLOR, M.Sc., JOHN CHONG, M.D., ROBIN ROBERTS, M.Sc., AND STEWART PUGSLEY, M.D.

Abstract Although the treatment of an individual patient in routine clinical practice has been likened to an experiment, the method is so susceptible to bias that we have come to demand multi-patient, double-blind, randomized controlled trials on matters of efficacy. Unfortunately, such trials have not or cannot be carried out for many clinical disorders; even when they have been executed their results may be difficult to extrapolate to individual patients.

To resolve this problem, we have begun to use double-blind randomized trials in which a single patient undergoes a series of pairs of treatments, consisting of one active and one placebo or alternative treatment per pair, with the order determined by random allocation. Appropriate treatment targets (signs, symptoms, or laboratory tests) are used as the measure of efficacy, and the trial is continued until efficacy is established or disproved. We describe such a trial, which resulted in a dramatically beneficial modification of treatment in a patient with partially reversible airflow limitation. We have established a clinical service that facilitates the widespread use of the method in our community. (N Engl J Med 1986; 314:889-92.)

SPECIAL ARTICLE

DETERMINING OPTIMAL THERAPY — RANDOMIZED TRIALS IN INDIVIDUAL PATIENTS

GORDON GUYATT, M.D., DAVID SACKETT, M.D., D. WAYNE TAYLOR, M.Sc., JOHN CHONG, M.D., ROBIN ROBERTS, M.Sc., AND STEWART PUGSLEY, M.D.

Abstract Although the treatment of an individual patient in routine clinical practice has been likened to an experiment, the method is so susceptible to bias that we have come to demand multi-patient, double-blind, randomized controlled trials on matters of efficacy. Unfortunately, such trials have not or cannot be carried out for many clinical disorders; even when they have been executed their results may be difficult to extrapolate to individual patients.

To resolve this problem, we have begun to use double-blind randomized trials in which a single patient undergoes a series of pairs of treatments, consisting of one active and one placebo or alternative treatment per pair, with the order determined by random allocation. Appropriate treatment targets (signs, symptoms, or laboratory tests) are used as the measure of efficacy, and the trial is continued until efficacy is established or disproved. We describe such a trial, which resulted in a dramatically beneficial modification of treatment in a patient with partially reversible airflow limitation. We have established a clinical service that facilitates the widespread use of the method in our community. (N Engl J Med 1986; 314:889-92.)

Menu

1. The Guyatt et al. (1986) example
2. What is an N-of-1 trial? *Definition*
3. Importance for health and life sciences: In general and for rare diseases in particular
4. Validity and methodological quality of N-of-1 RCTs
5. Data analysis in N-of-1 RCTs

1. The Guyatt et al. (1986) example

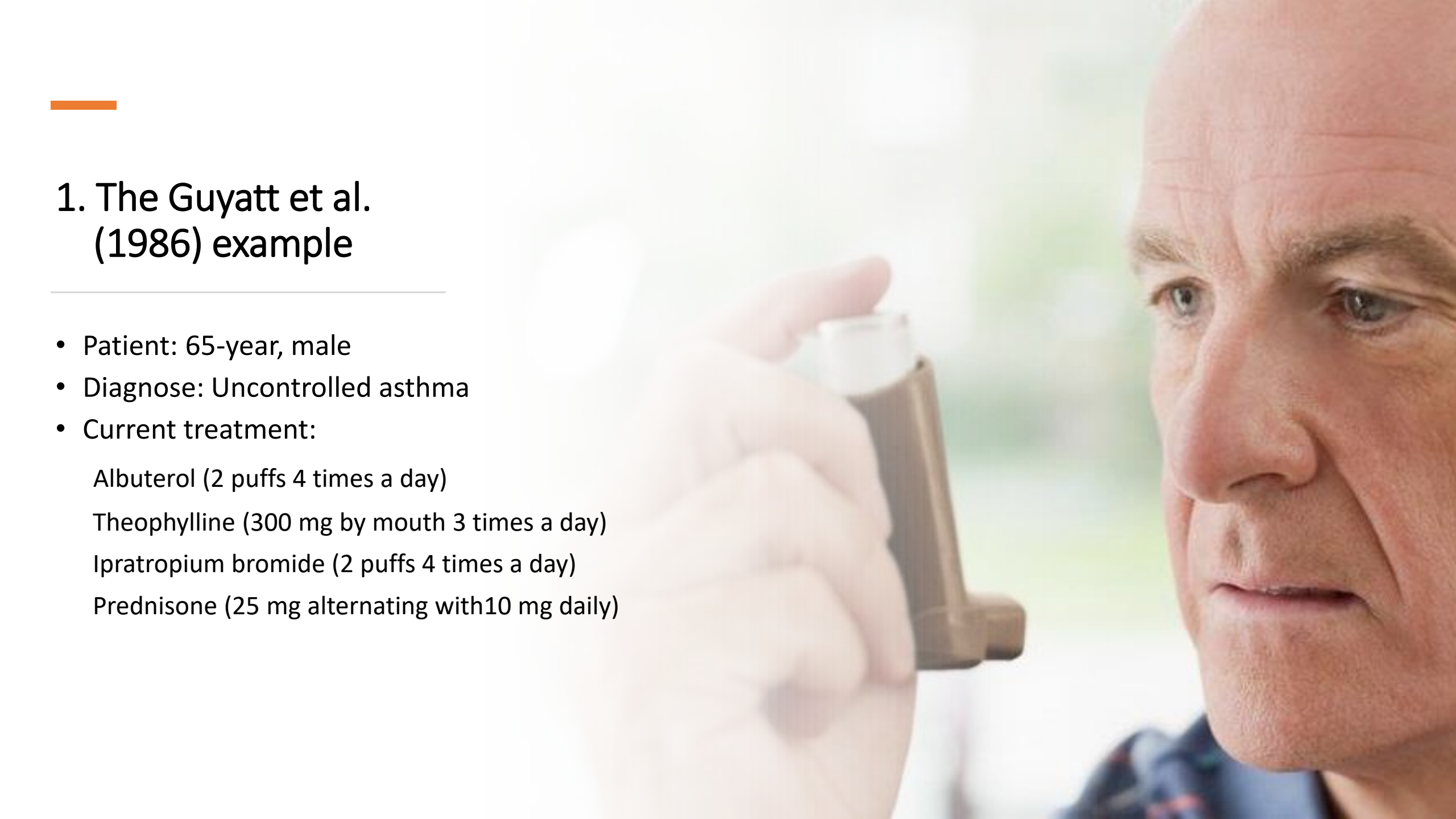
- Patient: 65-year, male
- Diagnose: Uncontrolled asthma
- Current treatment:

  Albuterol (2 puffs 4 times a day)

  Theophylline (300 mg by mouth 3 times a day)

  Ipratropium bromide (2 puffs 4 times a day)

  Prednisone (25 mg alternating with10 mg daily)



1. The Guyatt et al. (1986) example

- Patient: 65-year, male
- Diagnose: Uncontrolled asthma
- Current treatment:

  Albuterol (2 puffs 4 times a day)

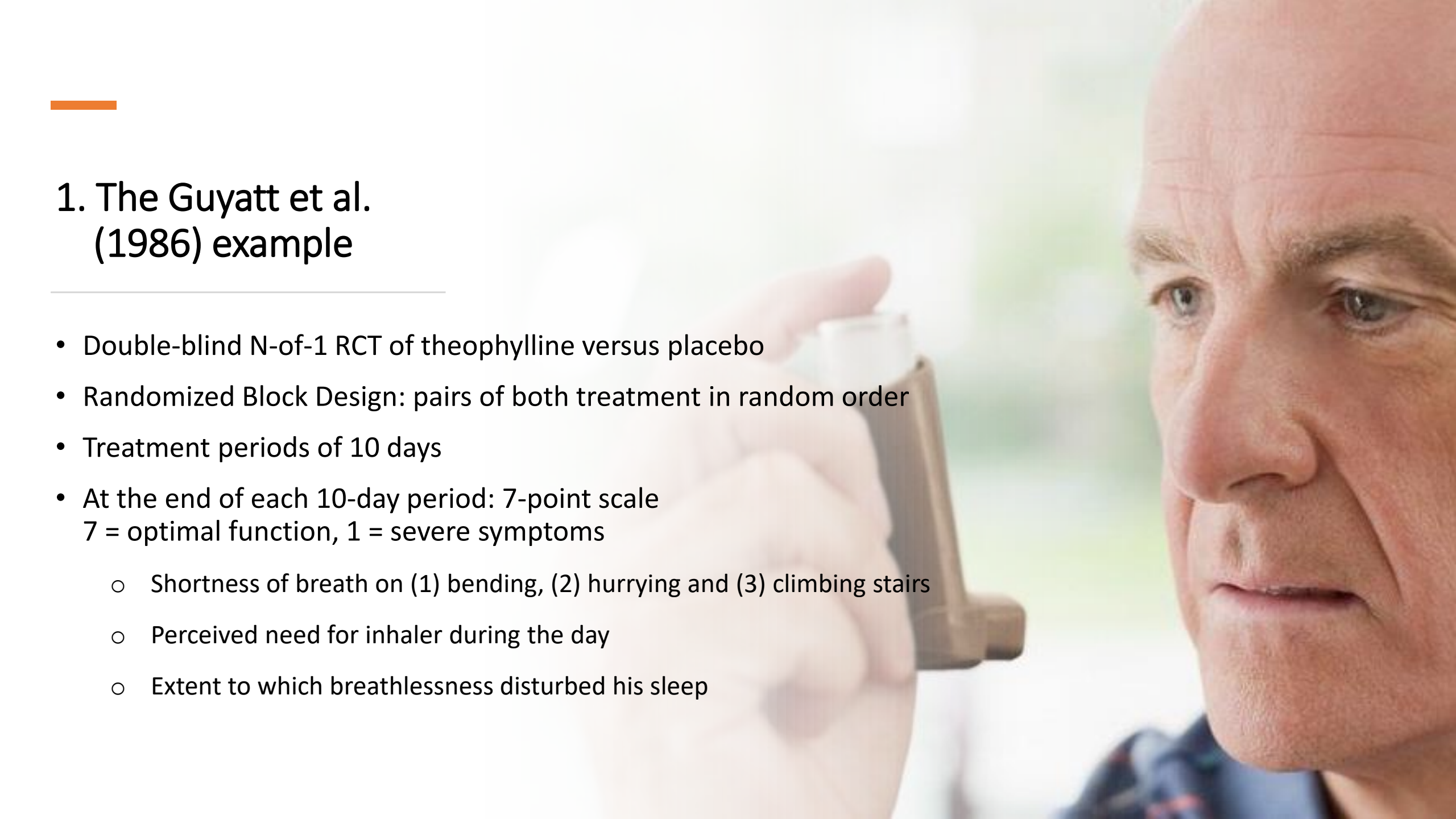
  Theophylline (300 mg by mouth 3 times a day) ?

  Ipratropium bromide (2 puffs 4 times a day) ????

  Prednisone (25 mg alternating with10 mg daily)

1. The Guyatt et al. (1986) example

- Double-blind N-of-1 RCT of theophylline versus placebo
- Randomized Block Design: pairs of both treatment in random order
- Treatment periods of 10 days
- At the end of each 10-day period: 7-point scale
  7 = optimal function, 1 = severe symptoms
  - Shortness of breath on (1) bending, (2) hurrying and (3) climbing stairs
  - Perceived need for inhaler during the day
  - Extent to which breathlessness disturbed his sleep



		*score**		
Shortness of breath	3	6	3	6
	3	5	3	5
	4	7	4	5
Need for inhaler	3	5.5	3	5
Sleep disturbance	5	5.5	3	5

*The patient rated his symptoms on a 7-point scale in which 7 represented optimal function and 1 represented severe symptoms.

Table 1. An N of 1 Randomized Controlled Trial of Theophylline.

SYMPTOM	PAIR 1		PAIR 2	
	PERIOD 1 (DRUG)	PERIOD 2 (PLACEBO)	PERIOD 1 (DRUG)	PERIOD 2 (PLACEBO)
		*score**		
Shortness of breath	3	6	3	6
	3	5	3	5
	4	7	4	5
Need for inhaler	3	5.5	3	5
Sleep disturbance	5	5.5	3	5

*The patient rated his symptoms on a 7-point scale in which 7 represented optimal function and 1 represented severe symptoms.

1. The Guyatt et al. (1986) example

- Double-blind N-of-1 RCT of ipratropium versus placebo
- Randomized Block Design: pairs of both treatment in random order
- Treatment periods of 10 days
- At the end of each 10-day period: 7-point scale
  7 = optimal function, 1 = severe symptoms
  - Shortness of breath on (1) bending, (2) hurrying and (3) climbing stairs
  - Perceived need for inhaler during the day
  - Extent to which breathlessness disturbed his sleep
- 3 repeated ratings during each period

	score											
Shortness of breath	6	6	6	6	6	6	6	6	6	5	5	5
	6	6	6	6	6	6	6	5	5	5	5	5
	4	5	4	5	5	5	5	4	5	4	5	5
Need for inhaler	4	4	4	5	5	5	5	5	5	4	4	4
Sleep disturbance	6	5	5	5	7	7	7	7	7	4	4	4

Table 2. An N of 1 Randomized Controlled Trial of Ipratropium.

Symptom	Pair 1						Pair 2					
	Period 1 (placebo)			Period 2 (drug)			Period 1 (drug)			Period 2 (placebo)		
	score											
Shortness of breath	6	6	6	6	6	6	6	6	6	5	5	5
	6	6	6	6	6	6	6	5	5	5	5	5
	4	5	4	5	5	5	5	4	5	4	5	5
Need for inhaler	4	4	4	5	5	5	5	5	5	4	4	4
Sleep disturbance	6	5	5	5	7	7	7	7	7	4	4	4

2. What is an N-of-1 trial? *Definition*

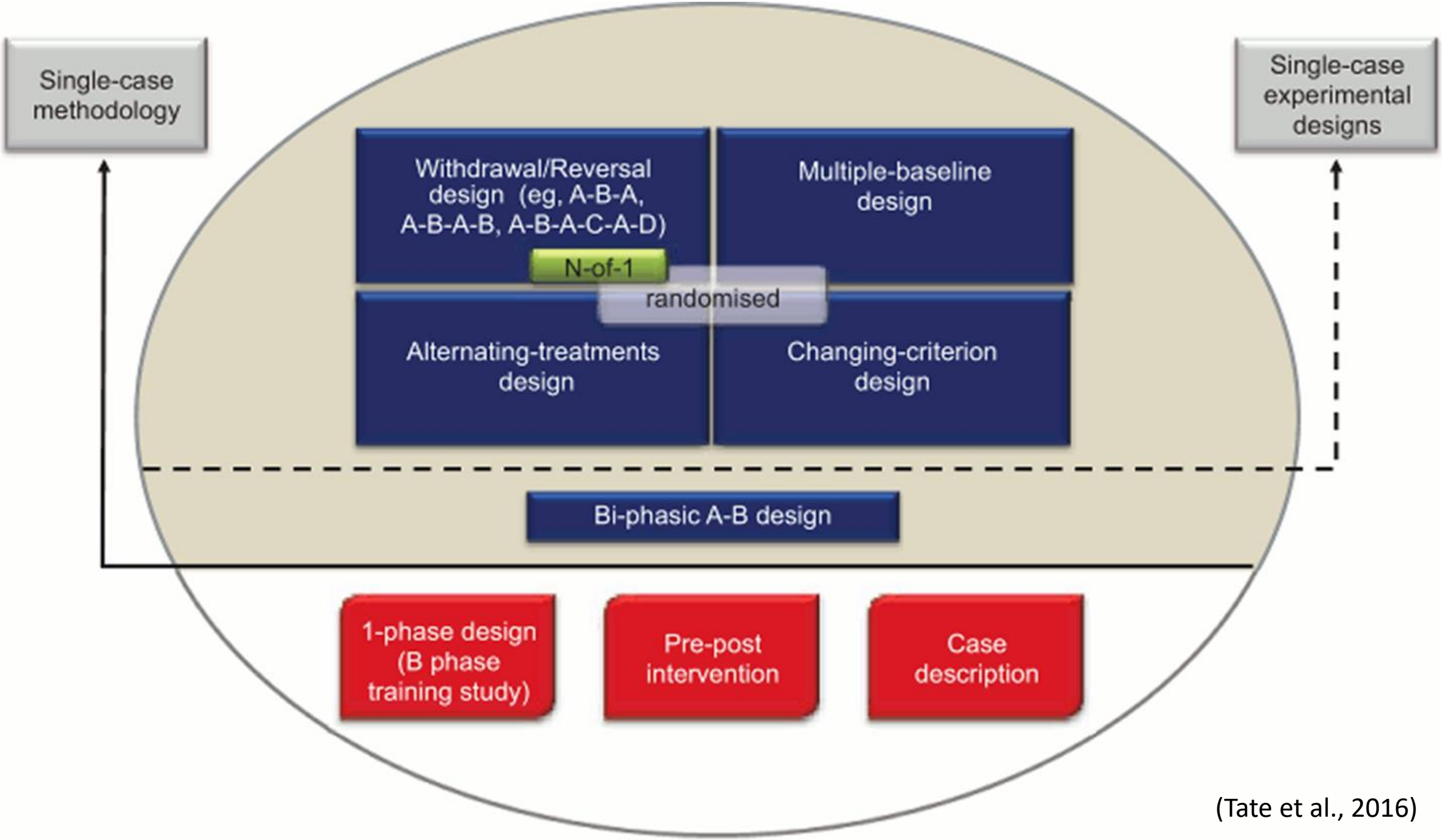
N-of-1 trial = a prospective, multiple crossover trial in a single patient

N-of-1 RCT = N-of-1 trial + randomization of the treatment sequence

Replicated N-of-1 RCTs = N-of-1 RCTs + replication across patients

≠ Case studies, case series, case reports, observational time series studies

= A specific design in a broader family of single-case experimental designs



(Tate et al., 2016)

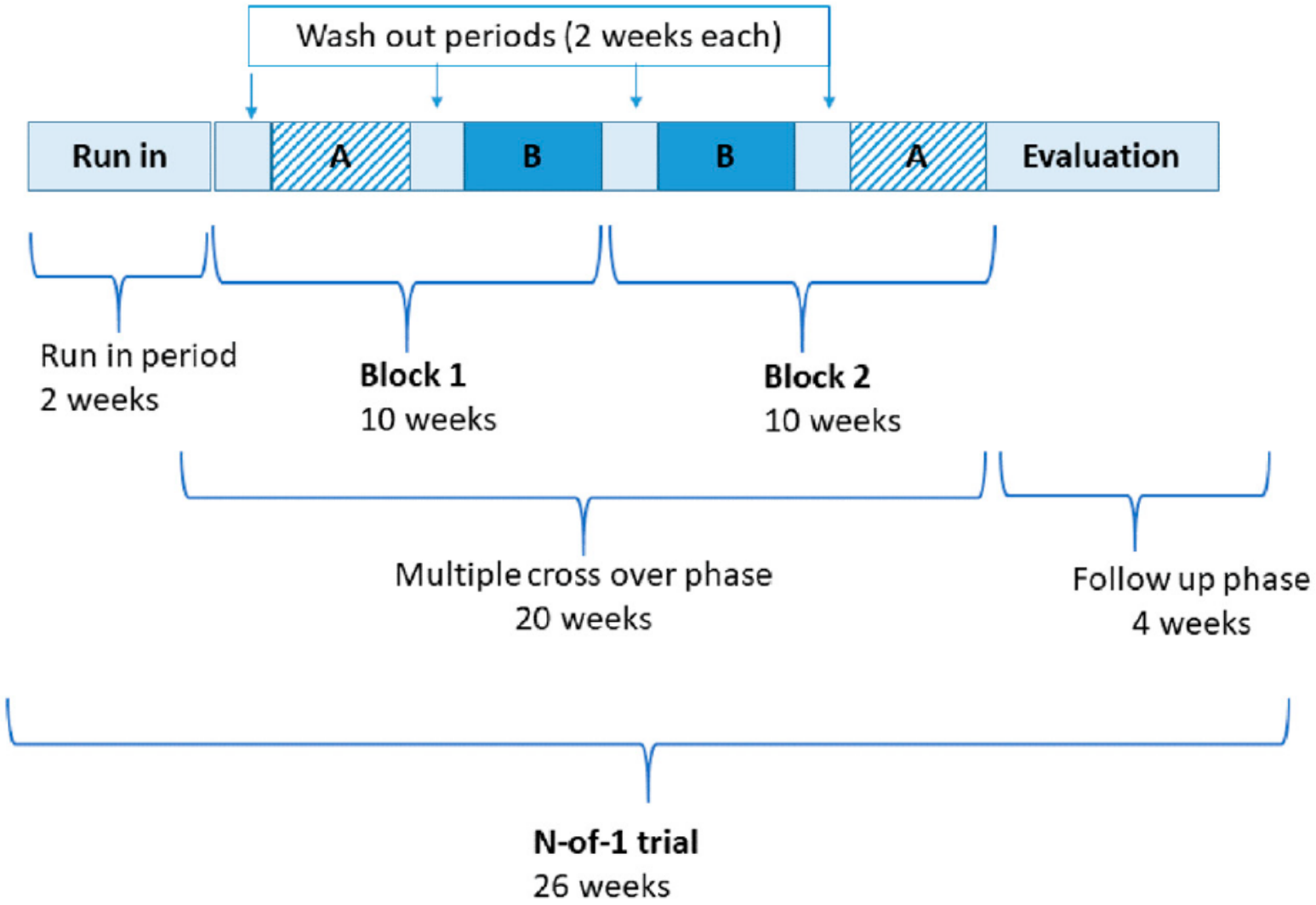


Figure 1. Simplified design of probiotics in fibromyalgia N-of-1 trial with timeline. A—denotes active supplementation; B—denotes placebo supplementation.

(Bradbury et al., 2020)

3. Importance for health and life sciences: In general and for rare diseases in particular

→ RCTs can answer the clinical research question: What works?

→ Large-scale group-comparison RCTs: What works **on average**?

≠ What works **in general**?

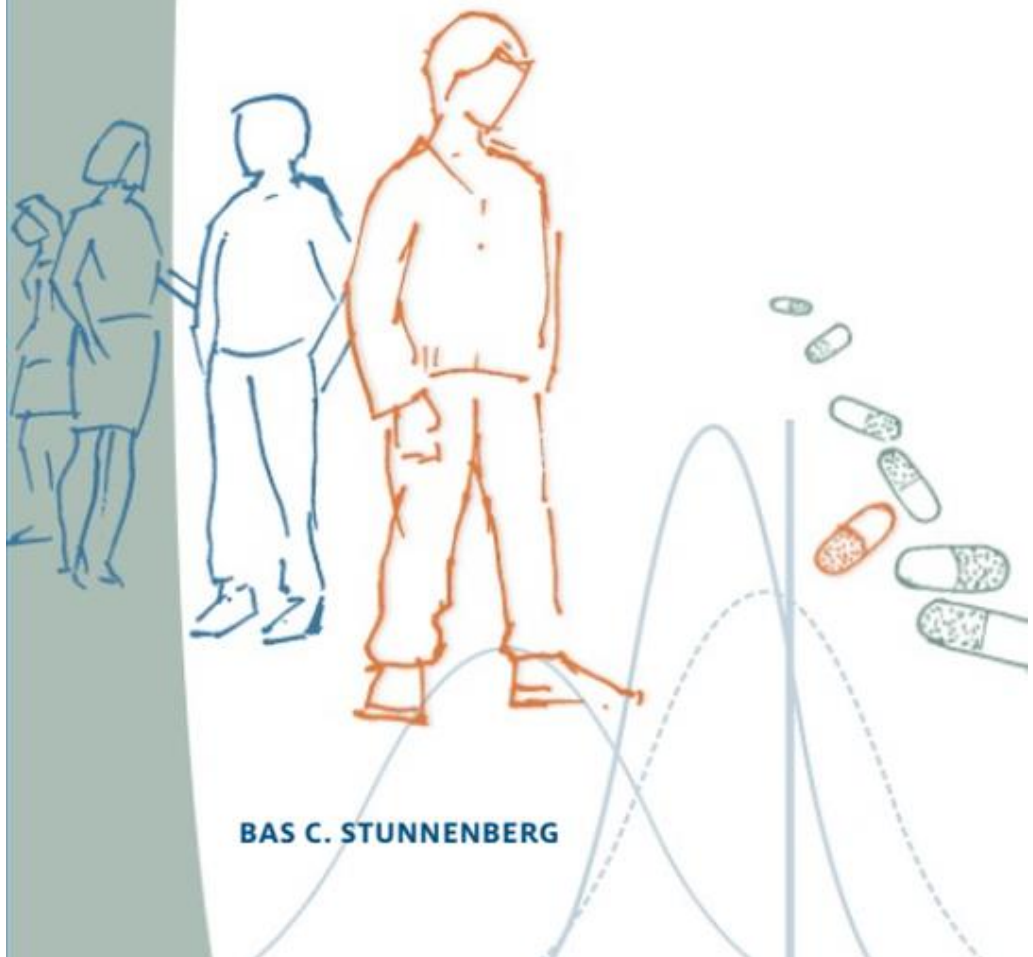
≠ What works **for the majority of patients**?

→ N-of-1 RCTs can answer the clinical research question:

What works **for this particular patient**?

N-of-1 trials for personalized treatment

THE CASE OF MUSCLE CHANNELOPATHIES



BAS C. STUNNENBERG

Building an evidence-base for the treatment of rare diseases

Rare diseases constitute a heterogeneous group of over 6.000 disorders with a prevalence of <1 per 2.000 per disease. In Europe, 30 million patients (6 to 8% of the population) are affected with one of these rare diseases.[1] Since neurological symptoms are present in about 75% of rare diseases[2], neurologist are familiar with the difficulties in determining the optimal therapy in patients with a rare disease while facing the paradox of evidence-based medicine (EBM) (see textbox 1).[3]

International regulatory authorities such as the Food and Drug Administration (FDA) and European Medical Agency (EMA) accept that it is unreasonable to demand the standard level of evidence (level 1) of multiple Randomized Controlled Trials (RCTs) in building an evidence-base for treatment of rare diseases.[4-6] The ability to conduct RCTs in rare diseases is hampered by low numbers of patients and large clinical heterogeneity. However, relying simply on case reports or case series incurs a considerable risk of selection and ascertainment bias. Currently, it is unclear which concessions can be accepted towards the level 1 evidence needed for registration and coverage decisions in case of rare diseases.[7,8]

Textbox 1 | The paradox of evidence-based medicine (EBM)

The paradox of EBM, first described by Guyatt et al. in 1986 in the New England Journal of Medicine[3], describes how physicians struggle to determine the optimal therapy for an individual patient in a 'scientific' fashion while dealing with the evidence-gap between clinical care and science: "Physicians cannot trust their own 'uncontrolled' therapeutic trials, but neither can they often look to large-scale randomized trials for definitive treatment recommendation". Guyatt and colleagues explain that, on the one hand, the uncontrolled therapeutic trials in clinical practice (where a treatment is provided for a certain time while the effect is estimated based on the subjective recollection of a response by the patient, sometimes supported by changes in physical examination or ancillary tests) are attributable to all kinds of sources of bias, such as disease fluctuations and the placebo (or nocebo) effect. But, on the other hand, the RCT, that deals with these important sources of bias by introduction of randomization and a placebo, only provides evidence on the effectiveness of a drug on a population level, i.e. an estimate of treatment effectiveness for the fictive 'average' patient. For numerous reasons, extrapolation of RCT trial results to inform treatment-decisions in an individual patient in clinical practice can be inappropriate (e.g. often the patient does not match the trials' inclusion criteria because of complex co-morbidity or co-medication). Finally, negative result from an RCT do not discard the possibility of some patients actually benefiting from the treatment (or vice versa).

VIEWS & REVIEWS OPEN ACCESS

Systematic Review of N-of-1 Studies in Rare Genetic Neurodevelopmental Disorders

The Power of 1

Annelieke R. Müller, MSc, Marion M.M.G. Brands, MD, PhD, Peter M. van de Ven, PhD, Kit C.B. Roes, PhD, Martina C. Cornel, MD, PhD, Clara D.M. van Karnebeek, MD, PhD, Frits A. Wijburg, MD, PhD, Joost G. Daams, MA, Erik Boot, MD, PhD, and Agnies M. van Eeghen, MD, PhD

Correspondence
Dr. van Eeghen
a.m.vaneeghen@amsterdamumc.nl

Neurology® 2021;96:529-540. doi:10.1212/WNL.0000000000011597

4. Validity and methodological quality of N-of-1 RCTs

BMJ 2008;337:a1655 doi: 10.1136/bmj.a1655 (Published 29 September 2008)

RESEARCH METHODS & REPORTING

N of 1 designs—Conventional trials aim to estimate the average effect of an intervention in a population. N of 1 trials, in which individuals undergo interventions with the order or scheduling decided at random, can be used to assess between and within person change and to investigate theoretically predicted mediators of that change

Developing and evaluating complex interventions: the new Medical Research Council guidance

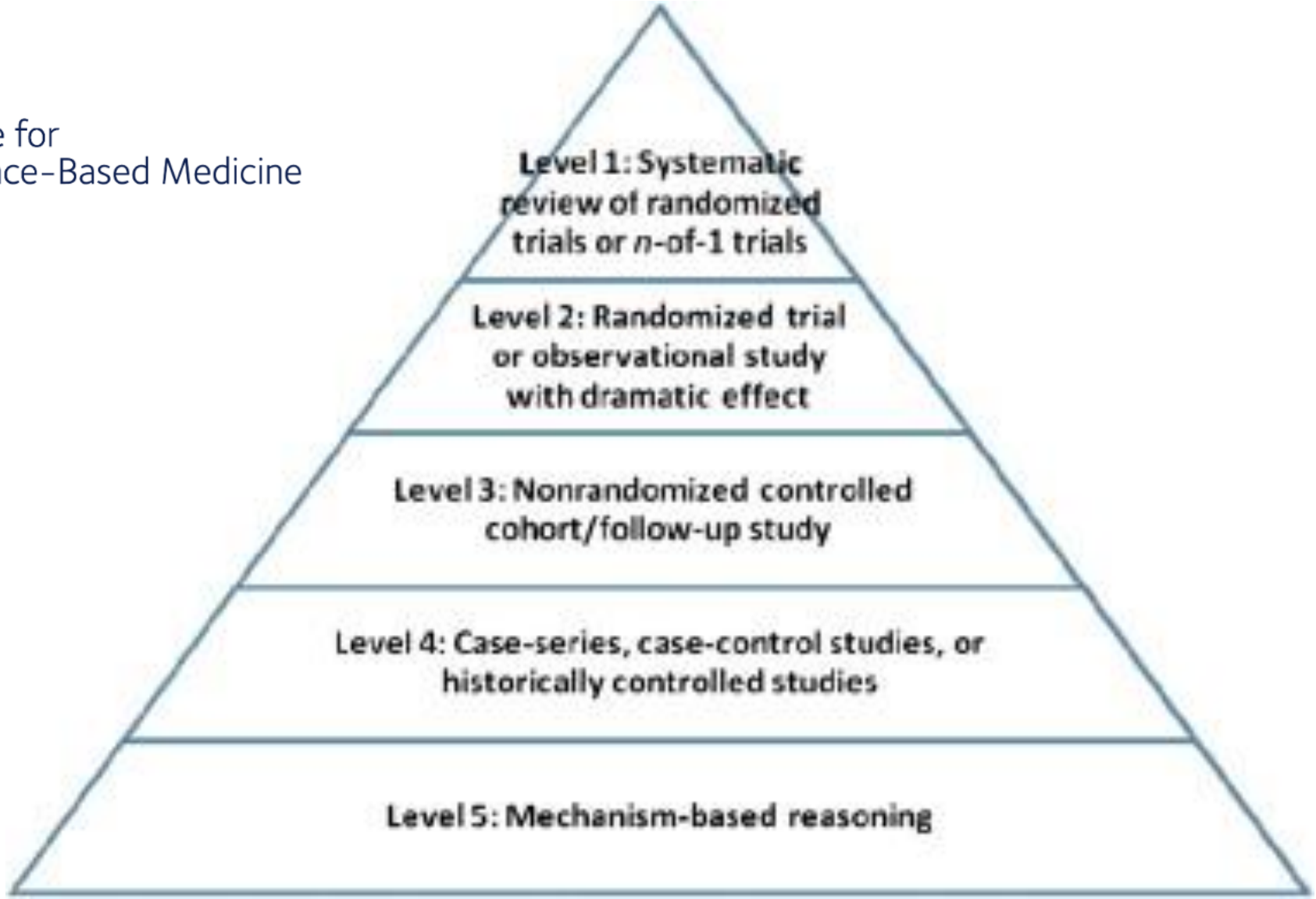
Evaluating complex interventions is complicated. The Medical Research Council's evaluation framework (2000) brought welcome clarity to the task. Now the council has updated its guidance

Peter Craig *programme manager*[1], Paul Dieppe *professor*[2], Sally Macintyre *director*[3], Susan Michie *professor*[4], Irwin Nazareth *director*[5], Mark Petticrew *professor*[6]

UNIVERSITY OF OXFORD

CEBM Centre for Evidence-Based Medicine

Oxford Centre for Evidence-Based Medicine 2011 Levels of Evidence



ICN MISSION STATEMENT AND OBJECTIVES

Vision: A world where personalised clinical studies (for both single and groups of individuals) are an integral part of clinical practice and health research

Mission: To promote, support and advance the use of personalised clinical studies, and to share relevant knowledge, experience, expertise, resources, and data through our global

KU LEUVEN

Symposium: Small is beautiful {once more}

Small is beautiful {once more}

The third international N=1 Symposium
April 2023

SMALL IS BEAUTIFUL
{ONCE MORE}

Within-person research has exponentially increased in recent years. Clinicians and researchers alike are benefiting from the flexibility and potential of research methods focusing on the individual. The result of such staggering interest has culminated in the current symposium, which strives to bring together experts and novices in this growing field of research. We aim to continue expanding awareness, knowledge, and expertise of this increasingly prominent research methodology.

Tweets

Tweets from @Nof1symposium

The 3rd International S...
@Nof1symposium · May 17

Replying to @Nof1symposium

... as @PatrickOnghena explained in his opening speech for the Small is Beautiful {once more} symposium #SCED

ELSEVIER

CrossMark

Journal of Clinical Epidemiology 76 (2016) 9–17

Journal of Clinical Epidemiology

CONSORT extension for reporting N-of-1 trials (CENT) 2015 Statement

Sunita Vohra[a,*], Larissa Shamseer[b], Margaret Sampson[c], Cecilia Bukutu[d], Christopher H. Schmid[e], Robyn Tate[f], Jane Nikles[g], Deborah R. Zucker[h], Richard Kravitz[i], Gordon Guyatt[j], Douglas G. Altman[k], David Moher[b], the CENT Group

SPIRIT extension and elaboration for n-of-1 trials: SPENT 2019 checklist

Antony J Porcino,[1] Larissa Shamseer,[2] An-Wen Chan,[3,4] Richard L Kravitz,[5] Aaron Orkin,[6,7] Salima Punja,[8] Philippe Ravaud,[9,10,11] Christopher H Schmid,[12] Sunita Vohra,[8,13] on behalf of the SPENT group

For numbered affiliations see end of the article.

Correspondence to: S Vohra svohra@ualberta.ca (ORCID 0000-0002-6210-7933)

Additional material is published online only. To view please visit the journal online.

Cite this as: *BMJ* 2020;368:m122 http://dx.doi.org/10.1136/bmj.m122

Accepted: 11 December 2019

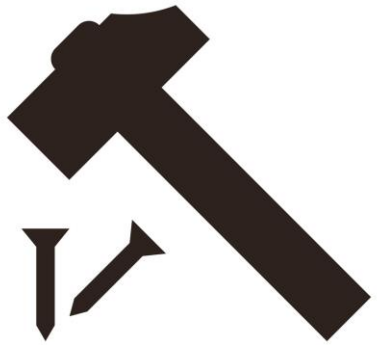
DOI: 10.1111/jebm.12484

ARTICLE

WILEY

The reporting quality of N-of-1 trials and protocols still needs improvement

Zhipeng Wei[1,2,3,4] | Xiajing Chu[1,2,3,4] | Jiani Han[1,2,3,4] | Na Zhang[1,2,3,4] | Yanfei Li[1,2,3,4] | Chaoqun Yang[1,2,3,4] | Qi Wang[5,6,7] | Jiang Li[1,4,8] | Ahmed Atef Belal[5,6,7] | Peijing Yan[9] | Xiuxia Li[1,2,3,4] | Kehu Yang[1,2,3,4]



5. Data analysis in N-of-1 RCTs

Graphical Data Analysis

Descriptive Statistics

Inferential Statistics

N-of-1 Randomized Trials

68

Reza D. Mirza, Sunita Vohra, Richard Kravitz, and Gordon H. Guyatt

S. Piantadosi, C. L. Meinert (eds.), *Principles and Practice of Clinical Trials*,
https://doi.org/10.1007/978-3-319-52636-2_97

Visual inspection alone: only for clinical use

***"The t-test is routinely used for N-of-1 RCTs, and is universally included in statistical packages."* (p. 1289)**

N-of-1 Randomized Trials

68

Reza D. Mirza, Sunita Vohra, Richard Kravitz, and Gordon H. Guyatt

S. Piantadosi, C. L. Meinert (eds.), *Principles and Practice of Clinical Trials*,
https://doi.org/10.1007/978-3-319-52636-2_97

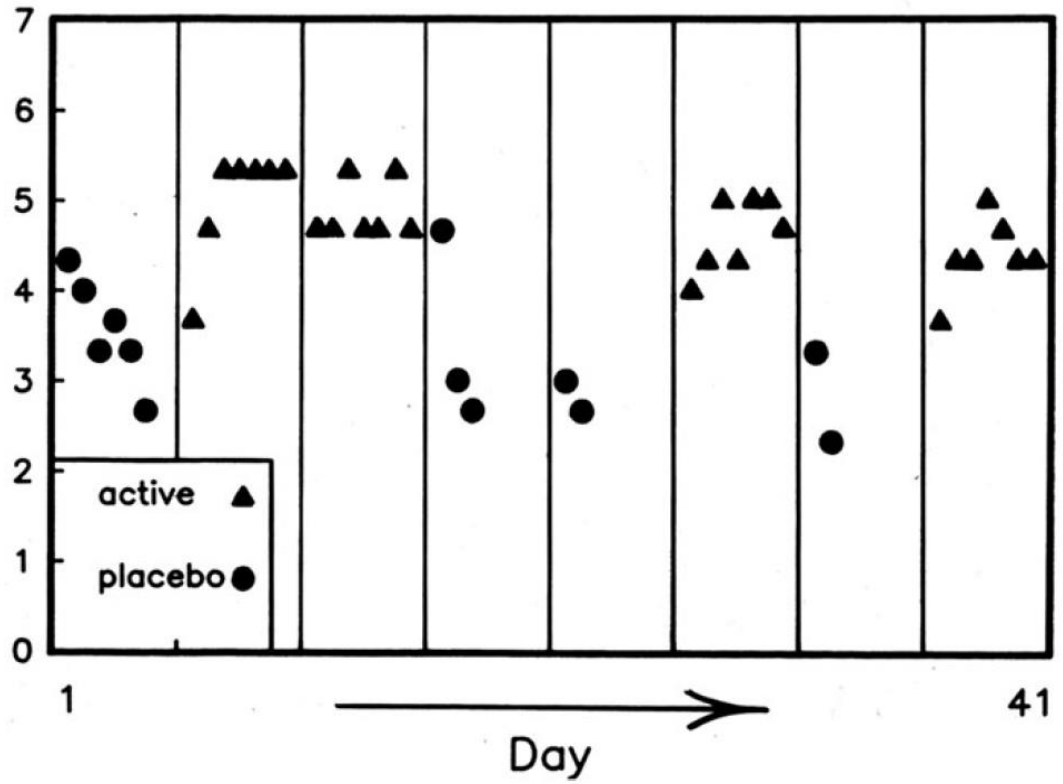


Fig. 1 N-of-1 RCT results: Mean daily Likert score

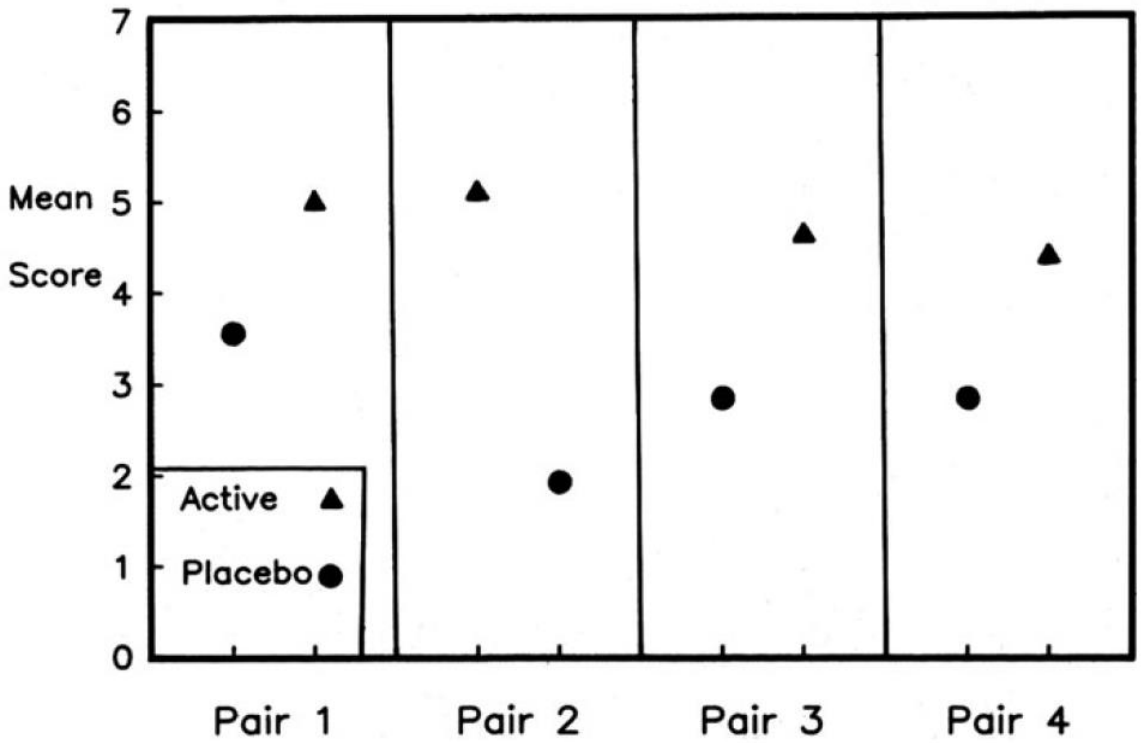


Fig. 2 N-of-1 RCT mean period score

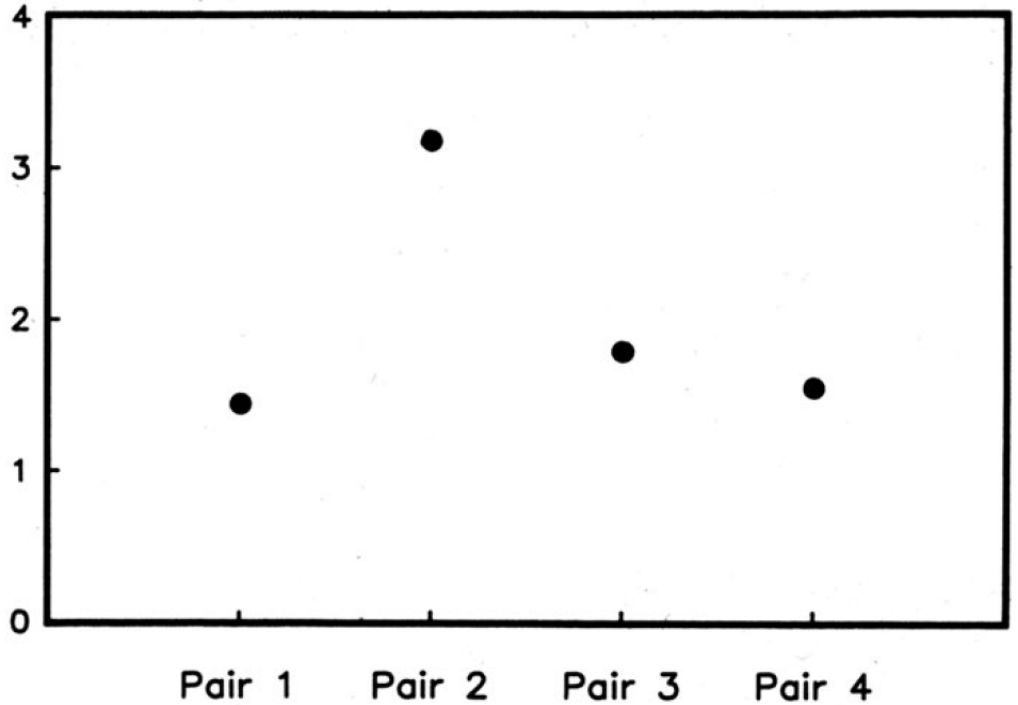


Fig. 3 N-of-1 RCT treatment and placebo difference scores

N of 1 RCT - Ms. A.D.

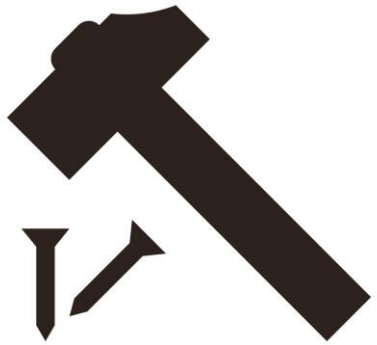
● Targets
(data--symptom means)

	Pair 1	Pair 2	Pair 3	Pair 4
Active	5.00	5.095	4.62	4.38
Placebo	3.56	1.98	2.83	2.83
Diff.	1.44	3.18	1.79	1.55

● Analysis (2 tailed paired t-test)

	Symptoms
$\bar{D}$	1.99
t	4.94
P	0.016
C.I. (90%)	(1.041, 2.937)

Fig. 4 N-of-1 RCT t-test results



Problems with the routine use of parametric *t*-tests in the analysis of N-of-1 RCT data

1. The pairs are not independent
2. The distributional assumptions of the test are implausible
3. The variability within a period is ignored
4. Missing data are ignored
5. Optional stopping requires additional Type I error rate control

Inferential data-analysis (Onghena et al., 2018, 2020)

- What is the statistical inference about?
  - Population = one particular patient
  - Sample = the repeated measures
  - Causal inference = demonstrations of a cause-and-effect relation for that specific patient
- Which statistical model?
  - Segmented linear and nonlinear regression models
  - Interrupted time series models – Borckard's Simulation Modeling Analysis
  - Multilevel models – Meta-analysis
- Which inferential procedure / logic?
  - Ordinary least squares and maximum likelihood criteria
  - Design-based – Randomization-based inference
  - Bayesian inference

Inferential data-analysis (Onghena et al., 2018, 2020)

- What is the statistical inference about?
  - Population = one particular patient
  - Sample = the repeated measures
  - Causal inference = demonstrations of a cause-and-effect relation for that specific patient
- Which statistical model?
  - Segmented linear and nonlinear regression models
  - Interrupted time series models – Borckard's Simulation Modeling Analysis
  - Multilevel models – Meta-analysis
- Which inferential procedure / logic?
  - Ordinary least squares and maximum likelihood criteria
  - **Design-based – Randomization-based inference**
  - Bayesian inference

N of 1 RCT - Ms. A.D.

- Targets
  (data--symptom means)

	Pair 1	Pair 2	Pair 3	Pair 4
Active	5.00	5.095	4.62	4.38
Placebo	3.56	1.98	2.83	2.83
Diff.	1.44	3.18	1.79	1.55

t = 5.10

B1 Active	5	A	A	A	A	A	A	A	P	P	P	P	P	P	P	P
B1 Placebo	3,56	P	P	P	P	P	P	P	A	A	A	A	A	A	A	A
B2 Placebo	1,98	A	A	A	A	P	P	P	P	P	P	P	A	A	A	A
B2 Active	5,095	P	P	P	P	A	A	A	A	A	A	A	P	P	P	P
B3 Active	4,62	A	A	P	P	P	P	A	P	A	P	A	A	P	A	P
B3 Placebo	2,83	P	P	A	A	A	A	P	A	P	A	P	P	A	P	A
B4 Active	4,38	A	P	P	A	A	P	P	A	A	P	P	P	P	A	A
B4 Placebo	2,83	P	A	A	P	P	A	A	P	P	A	A	A	A	P	P

B1	Active	5	A	A	A	A	A	A	A	P	P	P	P	P	P	P	P
B1	Placebo	3,56	P	P	P	P	P	P	P	A	A	A	A	A	A	A	A
B2	Placebo	1,98	A	A	A	A	P	P	P	P	P	P	P	A	A	A	A
B2	Active	5,095	P	P	P	P	A	A	A	A	A	A	A	P	P	P	P
B3	Active	4,62	A	A	P	P	P	P	A	P	A	P	A	A	P	A	P
B3	Placebo	2,83	P	P	A	A	A	A	P	A	P	A	P	P	A	P	A
B4	Active	4,38	A	P	P	A	A	P	P	A	A	P	P	P	P	A	A
B4	Placebo	2,83	P	A	A	P	P	A	A	P	P	A	A	A	A	P	P

t_OBS = 5.10

t1 t2 t3 t4 t5 t6 t7 t8 t9 t10 t11 t12 t13 t14 t15

$$P = \frac{\text{Number of test statistic values that are equal to, or more extreme, than the observed value}}{\text{Total number of test statistic values}}$$

$$P = \frac{2}{16}$$

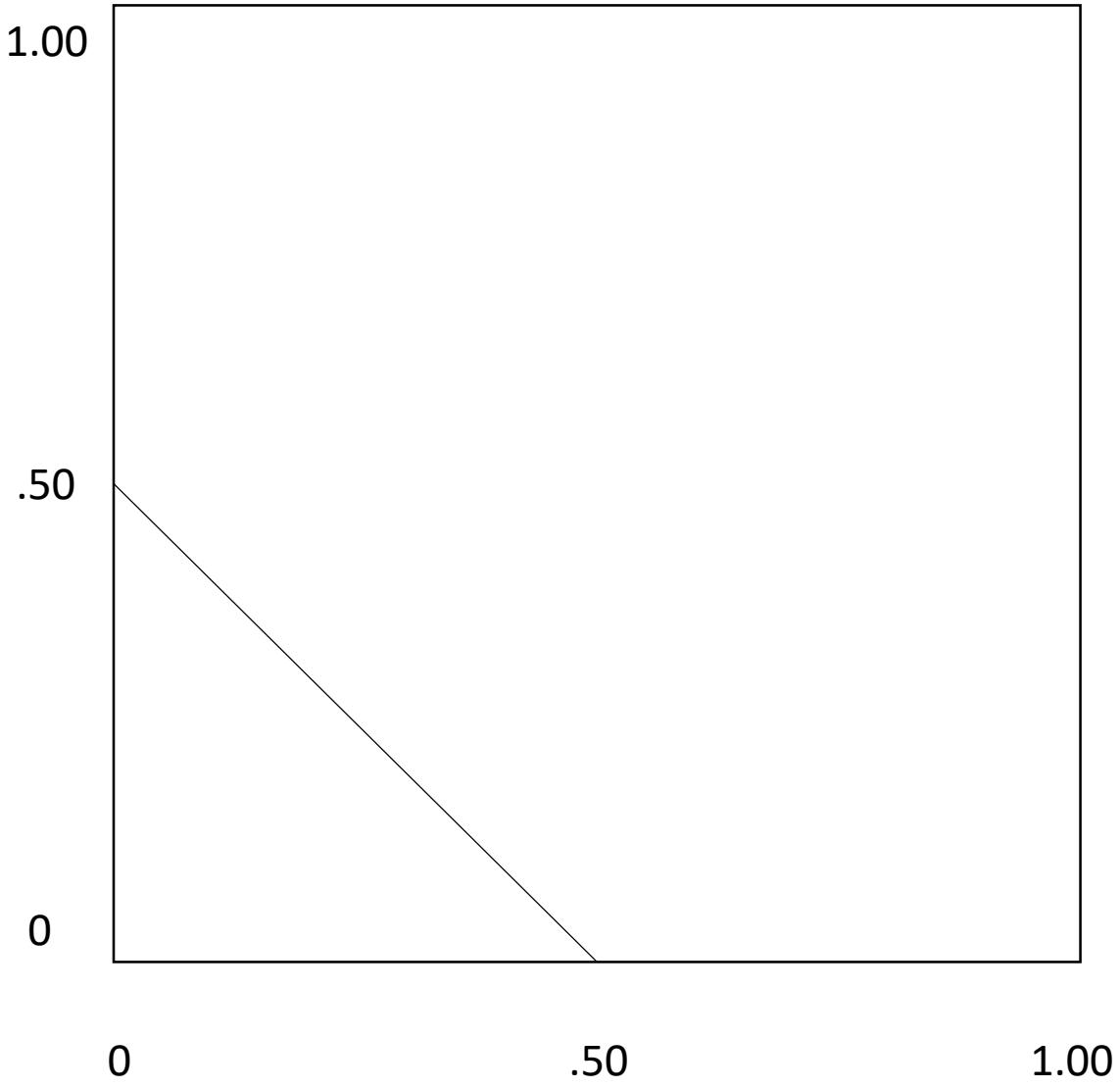
$$P = 0.125$$

```
> oneway_test(V3 ~ V2 | V1 , alternative='two.sided', distribution='exact',
+   data=Dataset)

        Exact Two-Sample Fisher-Pitman Permutation Test

data:  V3 by V2 (Active, Placebo)
        stratified by V1
Z = 1.8936, p-value = 0.125
alternative hypothesis: true mu is not equal to 0
```

Replications?



Example: $p_1 = .30$, $p_2 = .20 \rightarrow S = .50$

$P(S \leq .50)$?

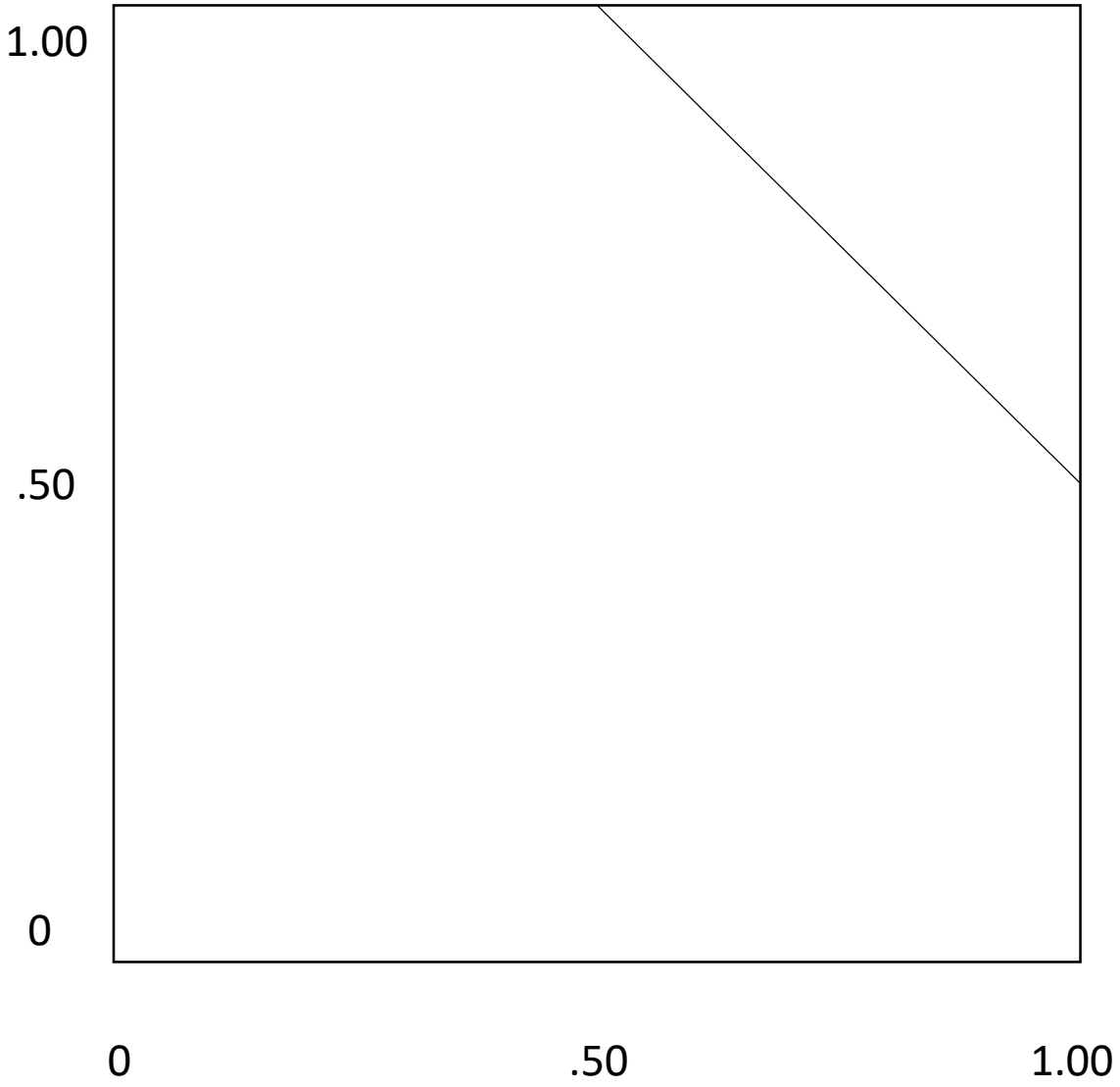
Under H_0: Uniform distribution

$P(S \leq .50) = (.50)^2/2 = .125$

$$P(S \leq S_{obs}) = \frac{(S_{obs})^2}{2}$$

(as long as the observed sum is not larger than 1)

Replications?



Example: $p_1 = .55$, $p_2 = .95 \rightarrow S = 1.50$

$P(S \leq 1.50)$?

Under H_0: Uniform distribution

$$P(S \leq 1.50) = \frac{(1.50)^2}{2} - (2)\frac{(0.50)^2}{2}$$

$$= 0.875$$

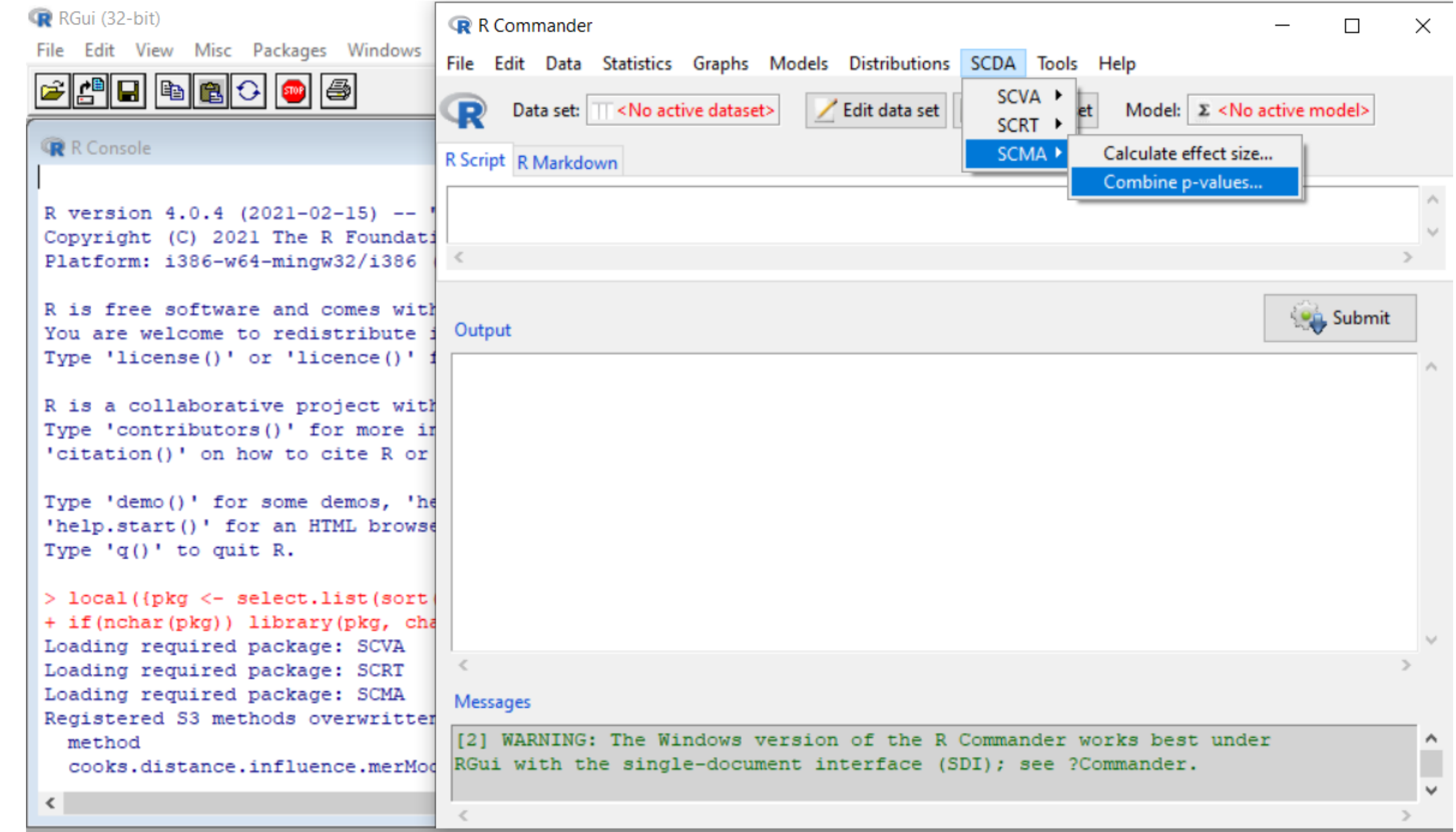
$$P(S \leq S_{obs}) = \frac{(S_{obs})^2}{2} - (S_{obs} - 1)^2$$

$$P(S \leq S_{obs}) = \sum_{k=0}^{\tilde{S}} (-1)^k \binom{n}{k} \frac{(S_{obs} - k)^n}{n!},$$

with n = the number of $P-$ values to be combined, and k = a counter up to the largest integer smaller than the observed sum $\tilde{S} = \max(k < S_{obs})$.

.30, .30, .20, .20 → .0417

(Edgington, 1972; Onghena & Edgington, 2005)


RGui (32-bit)
File Edit View Misc Packages Windows
R Console
R version 4.0.4 (2021-02-15) --
Copyright (C) 2021 The R Foundati
Platform: i386-w64-mingw32/i386
R is free software and comes with
You are welcome to redistribute i
Type 'license()' or 'licence()' f
R is a collaborative project with
Type 'contributors()' for more in
'citation()' on how to cite R or
Type 'demo()' for some demos, 'he
'help.start()' for an HTML browse
Type 'q()' to quit R.
> local({pkg <- select.list(sort(
+ if(nchar(pkg)) library(pkg, cha
Loading required package: SCVA
Loading required package: SCRT
Loading required package: SCMA
Registered S3 methods overwritten
method
cooks.distance.influence.merMod
R Commander
File Edit Data Statistics Graphs Models Distributions SCDA Tools Help
Data set: <No active dataset>
Edit data set
Model: <No active model>
SCVA
SCRT
SCMA
Calculate effect size...
Combine p-values...
R Script
R Markdown
Submit
Output
Messages
[2] WARNING: The Windows version of the R Commander works best under
RGui with the single-document interface (SDI); see ?Commander.

Plot observed data

Plot measure of central tendency

Plot estimate of variability

Plot estimate of trend

Plot interactive graph

Select the design type

ABAB Phase Design

X-axis label

Measurement Times

Y-axis label

Scores

A1 phase label

A1

B1 phase label

B1

A2 phase label

A2

B2 phase label

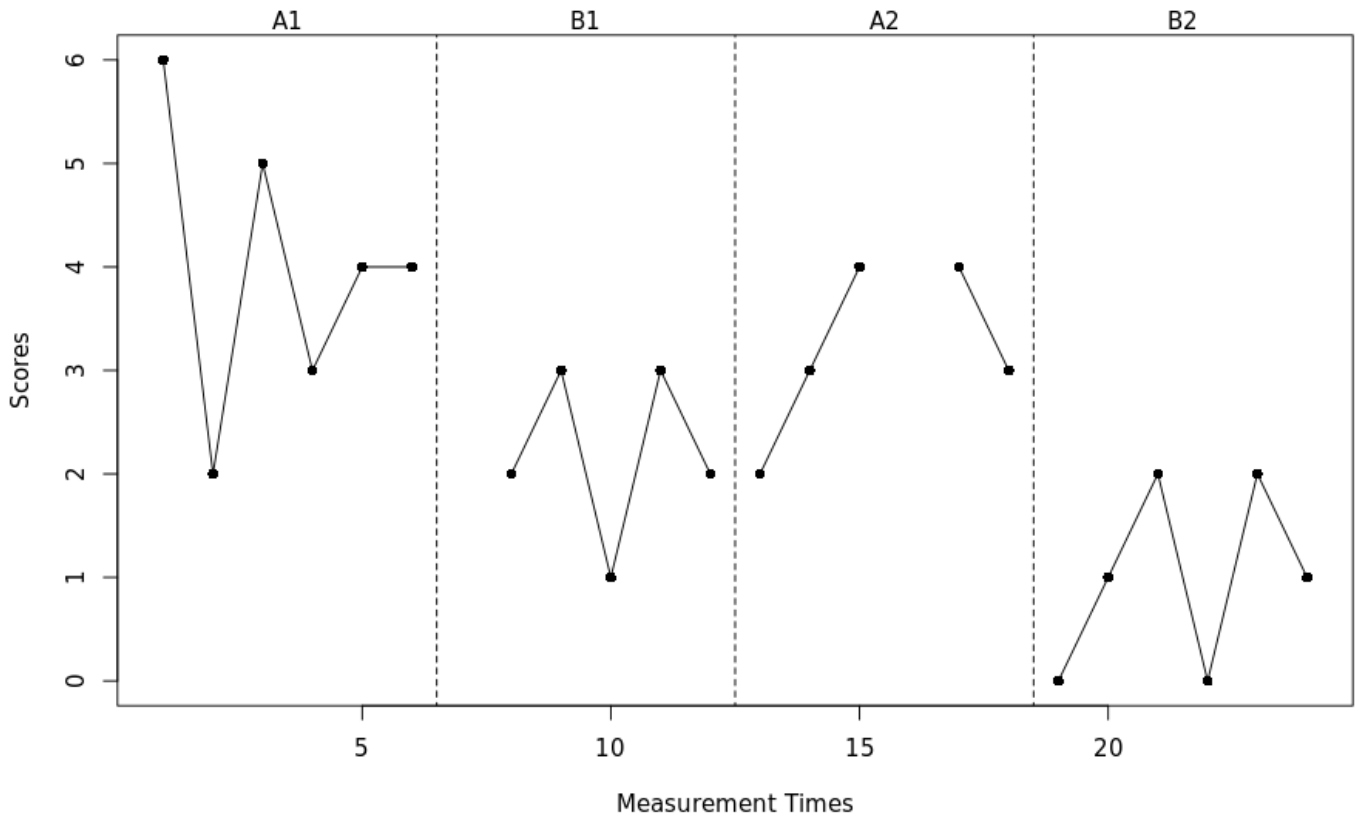
B2

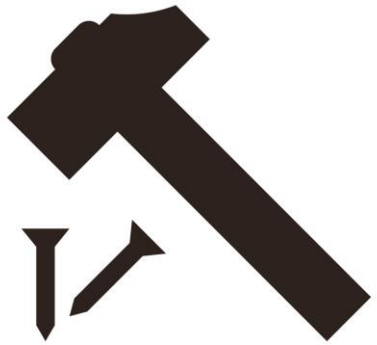
Y-axis minimum

Y-axis maximum

Plot

Plot





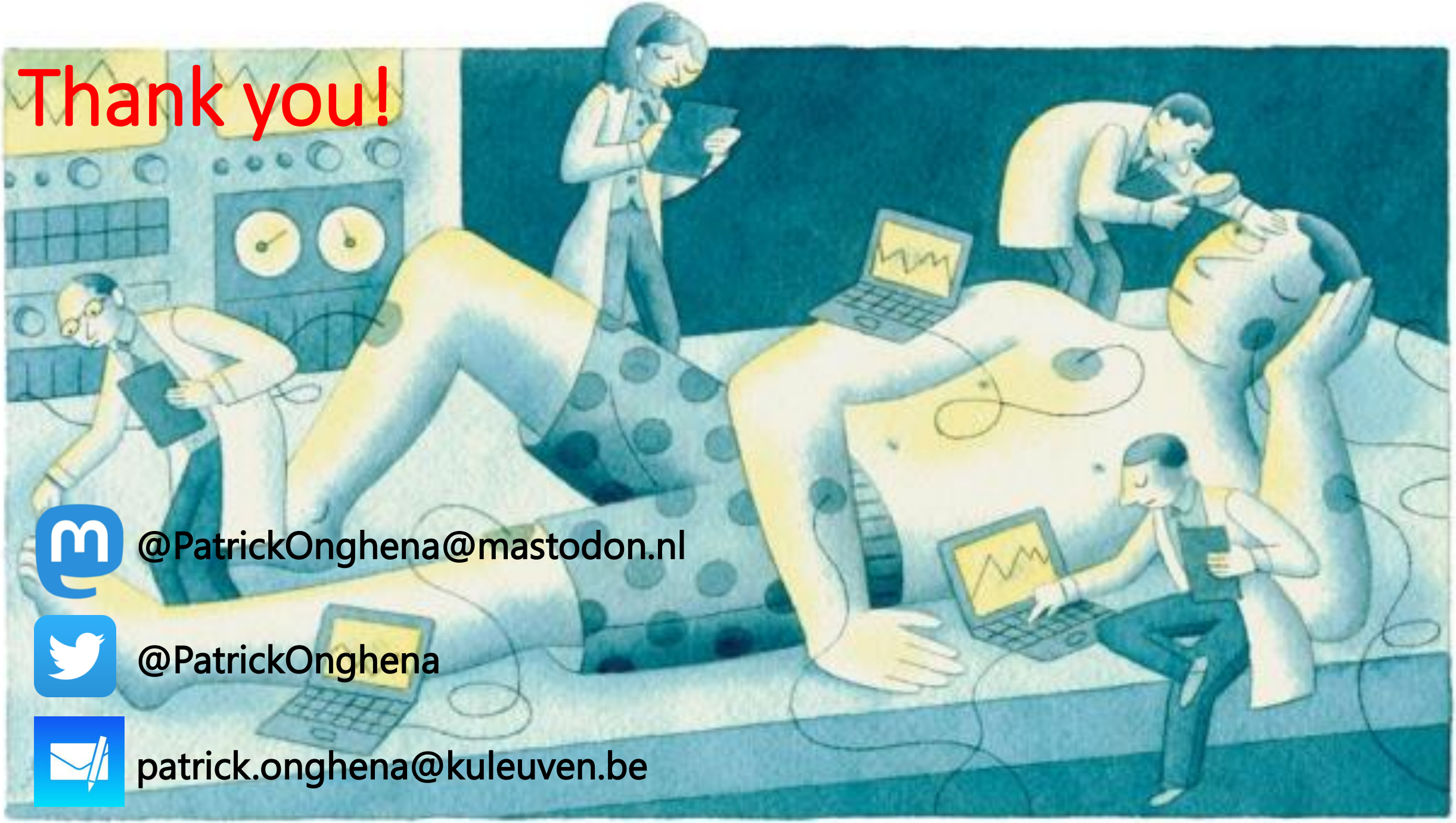
Shiny SCDA (De et al., 2020; De & Onghena, 2022)

1. The randomization tests do not assume independent data
2. The randomization tests are distribution-free
3. Variability within a period may be included by using other designs
4. Missing data are taken into account (even MNAR)
5. Optional stopping *not yet included*

Conclusion

1. Replicated N-of-1 RCTs have a long history, but only recently have been gaining popularity in the health sciences
2. Replicated N-of-1 RCTs are appealing for research on rare diseases because of their feasibility and because of their validity to test treatment effects at the individual level
3. Routine statistical analysis of N-of-1 RCT data needs to be improved
4. We need more user-friendly statistical tools and an effort in statistics education to move beyond the parametric t-test

Thank you!



@PatrickOnghena@mastodon.nl

@PatrickOnghena

patrick.onghena@kuleuven.be

References and further reading

Bradbury, J., Avila, C., & Grace, S. (2020). Practice-based research in complementary medicine: Could n-of-1 trials become the new gold standard? *Healthcare, 8*(1), 15. https://doi.org/10.3390/healthcare8010015

Bulté, I., & Onghena, P. (2008). An R package for single-case randomization tests. *Behavior Research Methods, 40*(2), 467–478. https://doi.org/10.3758/BRM.40.2.467

De, T. K., Michiels, B., Tanious, R., & Onghena, P. (2020). Handling missing data in randomization tests for single-case experiments: A simulation study. *Behavior Research Methods, 52*(3), 1355–1370. https://doi.org/10.3758/s13428-019-01320-3

De, T. K., & Onghena, P. (2022). The randomized marker method for single-case randomization tests: Handling data missing at random and data missing not at random. *Behavior Research Methods*, *54*(6), 2905–2938. https://doi.org/10.3758/s13428-021-01781-5

Edgington, E. S. (1972). An additive method for combining probability values from independent experiments. *The Journal of Psychology: Interdisciplinary and Applied, 80*(2), 351–363. https://doi.org/10.1080/00223980.1972.9924813

Edgington, E. S., & Onghena, P. (2007). *Randomization tests* (4th ed.). Chapman & Hall/CRC.

Guyatt, G., Sackett, D., Taylor, D. W., Chong, J., Roberts, R., & Pugsley, S. (1986). Determining optimal therapy: Randomized trials in individual patients. *The New England Journal of Medicine, 314*(14), 889–892. https://doi.org/10.1056/NEJM198604033141406

Michiels, B., Tanious, R., De, T.K., & Onghena, P. (2020). A randomization test wrapper for synthesizing single-case experiments using multilevel models: A Monte Carlo simulation study. *Behavior Research Methods, 52*(2), 654–666. https://doi.org/10.3758/s13428-019-01266-6

Nikles, J., & Mitchell, G. (2015). *The essential guide to n-of-1 trials in health.* Springer.

Nikles, J., Onghena, P., Vlaeyen, J., Wicksell, R., Simons, L.E., McGree, J.M., & McDonald, S. (2021). Establishment of an International Collaborative Network for N-of-1 Trials and Single-Case Designs. *Contemporary Clinical Trials Communications, 23,* Art.No. 100826, 1–8. https://10.1016/j.conctc.2021.100826

Onghena, P. (2020). One by one: The design and analysis of replicated randomized single-case experiments. In R. van de Schoot & M. Miočević (Eds.), *Small sample size solutions: A guide for applied researchers and practitioners* (pp. 87-101). Routledge. https://doi.org/10.4324/9780429273872

Onghena, P., & Edgington, E. S. (2005). Customization of pain treatments: Single-case design and analysis. *The Clinical Journal of Pain, 21*(1), 56–68. https://doi.org/10.1097/00002508-200501000-00007

Onghena, P., Michiels, B., Jamshidi, L., Moeyaert, M., & Van den Noortgate, W. (2018). One by one: Accumulating evidence by using meta-analytical procedures for single-case experiments. *Brain Impairment, 19*(1), 33–58. https://doi.org/10.1017/BrImp.2017.25

Porcino, A. J., Shamseer, L., Chan, A. W., Kravitz, R. L., Orkin, A., Punja, S., Ravaud, P., Schmid, C. H., Vohra, S., & SPENT group (2020). SPIRIT extension and elaboration for n-of-1 trials: SPENT 2019 checklist. *BMJ*, *368*, m122. https://doi.org/10.1136/bmj.m122

Tanious, R., De, T. K., & Onghena, P. (2019). A multiple randomization testing procedure for level, trend, variability, overlap, immediacy, and consistency in single-case phase designs. *Behaviour Research and Therapy, 119,* 103414. https://doi.org/10.1016/j.brat.2019.103414

Tate, R. L., Perdices, M., Rosenkoetter, U., Shadish, W., Vohra, S., Barlow, D. H., Horner, R., Kazdin, A., Kratochwill, T., McDonald, S., Sampson, M., Shamseer, L., Togher, L., Albin, R., Backman, C., Douglas, J., Evans, J. J., Gast, D., Manolov, R., Mitchell, G., … Wilson, B. (2016). The Single-Case Reporting Guideline In BEhavioural Interventions (SCRIBE) 2016 Statement. *Journal of Clinical Epidemiology*, *73*, 142–152. https://doi.org/10.1016/j.jclinepi.2016.04.006

Vlaeyen, J. W. S., Onghena, P., Vannest, K. J., & Kratochwill, T. R. (2022). Single-case experimental designs: Clinical research and practice. In G. J. G. Asmundson (Ed.), *Comprehensive clinical psychology: Volume 3* (2nd ed., pp. 1–28). Elsevier. https://doi.org/10.1016/B978-0-12-818697-8.00191-6

Vohra, S., Shamseer, L., Sampson, M., Bukutu, C., Schmid, C. H., Tate, R., Nikles, J., Zucker, D. R., Kravitz, R., Guyatt, G., Altman, D. G., Moher, D., & CENT Group (2015). CONSORT extension for reporting N-of-1 trials (CENT) 2015 Statement. *BMJ*, *350*, h1738. https://doi.org/10.1136/bmj.h1738