Trends in Medicine

Perspective Article ISSN: 1594-2848

Demystification – a solution for assessment of real-world effectiveness?

Franz Porzsolt[1]*, Christel Weiss[1,2], Manfred Weiss[1,3], Karl-Walter Jauch[1,4] and Robert M Kaplan[5]

[1]Institute of Clinical Economics (ICE) e.V., D-89081 Ulm/Germany

[2]Department Medical Statistics and Biomathematics, Univ. of Heidelberg, D-68617 Mannheim/Germany

[3]Department Anaestesiology, University of Ulm, D-89081 Ulm/Germany

[4]University Hospital, Ludwigs-Maximilians-University München, D-81377 München/Germany

[5]Clinical Excellence Research Center, University of Stanford, Stanford/CA 94305-6015, USA

Perspective in response to the article in NEJM & (sounding board): The Magic of Randomization versus the Myth of Real-World Evidence. by Collins R, Bowman L, Landray M, Peto R. N Engl J Med 2020;382:674-678

Collins and colleagues argue the magic of randomization can compensate for the myth of real-world conditions (RWC) [1]. This assertion is problematic for two reasons: randomization cannot be completed under RWC, but artificial intelligence may demystify the RWC.

Nearly 100 years ago Archie Cochrane and Austin Bradford Hill proposed a three-dimensional assessment for evaluating clinical interventions: Can it Work? Does it work? Is it Worth it? [2, S1]. Table 1 proposes a strategy to answer these questions by description of three outcome dimensions that are assessed under two different conditions from three different perspectives. The strategy requires a different architecture of studies (form and function) and different tools.

The difference between RWC and ESC

Three research groups [3,S2,S3] defined the allocation as rule for differentiation of experimental and pragmatic studies. The investigator allocates under ESC, the practitioner under RWC. Additional difference of ESC and RWC are shown in Appendix I. Exclusion criteria that eliminate participants with known risk factors under ESC, do not exist under RWC. Under RWC participants select themselves and physicians chose the best possible interventions for each individual patient. From the investigators perspective this day-to-day clinical practice appears like a 'Natural Chaos'.

Even randomized controlled trials (RCTs) have limitations

Although RCTs are considered the Gold Standard in clinical research several limitations are often overlooked. We consider four examples. The RCT is expected to distribute known and unknown confounders equally among study groups. This assumption is justified in case of 'adequate trial size'. We estimated the minimal necessary size of RCTs that can guarantee the equal distribution of confounders based on four assumptions: 10 confounders, independent from each other, dichotomous distribution, and maximal tolerated difference of 5%. The result of this model revealed a minimal number of 1000 participants per trial to guarantee the fairly equal distribution of confounders [S4, S5]. Most studies that support clinical recommendations include less than 1000 patients (Appendix II).

Second, open (unblinded; unmasked) studies are biased by the strong and weak expectations of both patients and attending practitioners. Unmet strong expectations result in refusal of participation and contribute to sampling bias. Patient expectations were classified as 'weak' if their expectations are not met, but the patient nevertheless participates in a study. It is shown in Appendix III that large difference in the expectations towards the alternative treatment arms (e.g. 80% of patients prefer intervention 'A' and 20% 'B') predict that the treatment group with the higher preference rate will achieve better study results [S6-S10]. Third, de-masking or de-blending during the course or the study may cause bias. Several examples demonstrate the effects of deblending [S11,S12]. Recently a meta-meta-analysis claimed that blinding has no effect on the results of RCTs [S13], yet a simple selection bias should be excluded, which may explain the reported effect [S14]. Fourth, non-inferiority designs can only be used to compare interventions with previously confirmed benefit [S15]. Non-inferiority is difficult to interpret unless we have evidence that both the experimental group and the comparator are superior to placebo [S16].

The RWC: a 'Natural Chaos' from the perspective of the investigator

In clinical practice, decisions are supported either by external evidence (based on published data), internal evidence (based on personal experience) or both. As an example of 'Natural Chaos', we compared 330 recommendations of international guidelines from 11 countries and observed congruence in recommendations i.e. at least 66% congruency in 50 of 330 (15%) recommendations, incongruence in 213 (65%), and undetermined recommendations in 67 (20%). One of the most likely interpretations is that clinical experts who review these guidelines are exposed to two different types of information. When treating patients, they observe effectiveness data under RWC. Scientific publications describe efficacy data generated under ESC. It is likely that both efficacy and effectiveness data influence the opinions of decision makers differently [S17- S19] and the different information may lead to the observed incongruence. Our 42 tables summarizing this literature are available upon request.

Correspondence to: Franz Porzsolt, Institute of Clinical Economics (ICE) e.V., D-89081 Ulm/Germany, E-mail: f.porzsolt@gmx.de

Received: April 18, 2020; **Accepted:** May 13, 2020; **Published:** May 15, 2020

Table 1. The tree-dimensional strategy according to Cochrane und Hill

Question	Outcome dimension	Condition of study	Per-spective	Architecture of study Form (design)	Function	Tool
Can it work?	**Efficacy** (Proof of Principle)	Experimental Study Condition (**ESC**)	Clinical Research	Explanatory or interventional study	Demonstration of Proof of principle **(PoP)**	Randomized Controlled Trial [**RCT**]
Does it work?	**Effectiveness** (Real World Effectiveness)	Real World Condition (**RWC**)	Health Services Research	Pragmatic or observational study	Confirmation of Real World Effectiveness **(RWE)**	Pragmatic Controlled Trial [**PCT**]
Is it worth it?	**Value** (Subjective individual and societal value)		Subjective individual & societal	Complete Economic Analysis	Comparison of costs & consequences of different actions **(Value)**	Cost-Effectiveness Analysis [**CEA**]

A former version of this table is published in: Porzsolt F, Weiss Ch, Weiss M, Müller AG, Becker SI, Eisemann M, Kaplan RM. Versorgungsforschung braucht dreidímensionale Standards zur Beschreibung von Gesundheitsleistungen [Health services research needs three-dimensional standards for description of health services]. Monitor Versorgungsforschung 2019;04:53-60. http://doi.org/10.24945/MVF.04.19.1866-0533.2163.

The solution of the unsolved problems

RCTs can address the "Can it work? question while neglecting evidence about effectiveness and value. The appropriate tool for assessment of RWE and Value, the Pragmatic Controlled Trial [4,5,S20-S22] is based on four principles. First, the practitioner and patient but not the investigator decide about the allocation. Second, the investigator defines which artificial intelligence tools should be applied to render the 'Natural Chaos' evaluable under RWC. Third, these tools, e.g. Bayes Theorem [S23] are used to stratify the real-world data for evaluation. Fourth, allocation and evaluation must be independent from each other.

This solution classifies all patients with a comparable clinical problem (e.g. breast cancer) who are treated individually according to their risks profiles each related to three different endpoints: main goals e.g. mortality, main treatment side effects e.g. allergic reactions, and total costs of treatment (Appendix IV). The risk profiles of the groups had to be predefined before start of the trial. All study participants are allocated to a high, intermediate or low risk group separately for each endpoint by using a separate algorithm for classification of the risk groups for each of the assessed endpoints. The steering group selects the interventions that will be compared with each other and with a mixed group of the remaining 'any other interventions'. Corrections for multiple testing are necessary.

All patients – without exception – who meet the inclusion criteria e.g. breast cancer and ask for help at a participating center within a defined time window are included in a Pragmatic Controlled Trial (PCT). This procedure meets four essential requirements: A) the practitioner and patient decide about the allocation as requested for pragmatic trials [3,S2,S3]. B) the investigator defines the rules that enable the unbiased evaluation of the 'Natural Chaos' according to the Bayes Theorem [S23]. C) patient self-selection defines the membership to the real-world patient population. D) the trial includes a well-defined control group i.e. the mixed but risk-stratified intervention groups that were not selected as specific intervention target groups.

Research progress and advantages

The goal of this strategy is to relate outcomes to baseline risks. It differentiates three outcome dimensions and reduces the number of human experiments to an absolute minimum. Two or three high quality RCTs will be sufficient for demonstration of the Proof of Principle, or "Can it work". Ethical conflicts associated with unnecessary human experiments, the high costs and complexity of conducting randomized trials, and bureaucratic burden will be reduced. We need bold new approaches to complete a paradigm shift in healthcare from a mono-dimensional to the three-dimensional assessment of outcomes that was suggested almost 100 years ago by Sir Archie Cochrane and Sir Austin Bradford Hill [2,S1].

Second, the new tool was used to supplement the missing part in the three-dimensional Cochrane-Hill strategy. The work resulted from the collaboration of a large group of American, Brazilian, Italian, Norwegian, and German colleagues and doctoral students. Table 1 summarizes the consensus from this collaboration (Supplementary references [S1] – [S23] are listed in Appendix V).

References

1. Collins R, Bowman L, Landray M, Peto R (2020) The magic of randomization versus the myth of real-world evidence. *N Engl J Med* 382: 674-678. [Crossref]
2. Haynes B (1999) Can it work? Does it work? Is it worth it? The testing of healthcare interventions is evolving. *BMJ* 319: 652-653. [Crossref]
3. Schwartz D, Lellouch J (1967) Explanatory and pragmatic attitudes in therapeutic trials. *J Chron Dis* 20: 637-648. [Crossref]
4. Porzsolt F, Eisemann M, Habs M, Wyer P (2013) Form follows function: Pragmatic controlled trials (PCTs) have to answer different questions and require different designs than randomized controlled trials (RCTs). *J Publ Health* 21: 307-313. [Crossref]
5. Porzsolt F, Rocha NG, Toledo-Arruda AC, Thomaz TG, Moraes C, et al. (2015) Efficacy and effectiveness trials have different goals, use different tools, and generate different messages. *Pragmatic and Observational Research* 6: 47-54. [Crossref]