RESEARCH ARTICLE

The Front-End Processor Developed by Engineers — A Useful Tool for Describing the Quality and Quantity of Progress in Healthcare

Franz Porzsolt, Meret Phlippen, Philipp Legrum, Manfred Weiss[1]

1 University Hospital of Ulm (Universitätsklinikum Ulm)

Funding: No specific funding was received for this work.

Potential competing interests: No potential competing interests to declare.

Abstract

Introduction. The ‘natural chaos’ of seemingly unstructured healthcare can be analyzed under structured conditions. This study describes methods for quantification of progress in healthcare. *Methods.* Engineers published the “front-end processor (FEP)” method to detect and eliminate defects in steel production. We use three scenarios to demonstrate that the FEP can confirm the congruence of expectations, existing data and derived results in healthcare. (1) Six teams analyzed the validity criteria in each of 20 publications to confirm the congruence of the initial study question with its mathematical confirmation. (2) Different strategies and methods answer the three Cochrane questions, i.e. Can it work?, Does it work? Is it worth it?. (3) Traditional 2 x 2 contingency table quantify two different sets of information, the traditional confirmation/exclusion of a suspected disease and the induction of its psychologic effects. *Results.* (1) Four steps were identified from the simple study question to the mathematical confirmation of the answer. (2) Two functional and twelve formal criteria characterize the experimental study condition and the two pragmatic conditions of care, either with or without systematic analysis of the results. Experiments use the randomized controlled trial (RCT), pragmatic controlled trials (PCTs) use the Bayesian statistics. (3) The communication of ‘bad news’ increases fear/anxiety in patients 10 – 5000 fold, but ‘good news’ has almost no effect on the perceived safety. *Discussion:* We confirm that the "FEP" developed by engineers can confirm the expected advances in healthcare and thus increase the validity of ethical, medical, epidemiological, economic, legal, and political consequences.

Franz Porzsolt[1,a,*], **Meret S. Phlippen**[1,2,b], **Philipp Legrum**[1,c], and **Manfred Weiss**[1,3,d]

[1] *Private Research Institute Clinical Economics (PRICE; Ringgold iD: 587744), 89081 Ulm / Germany*

[2] *Vivantes Clinic in Friedrichshain Berlin, HNO Clinic, 10249 Berlin / Germany*

[3] *Clinic for Anaesthesiology and Intensive Medicine, Ulm University Clinic, 89081 Ulm / Germany*

[a] ORCID iD: 0000-0003-3554-2902

[b] ORCID iD: 0000-0003-0423-4448

[c] ORCID iD: 0000-0002-7548-7277

[d] ORCID iD: 0000-0003-4851-6439

[*]Corresponding author:
Franz Porzsolt MD, PhD, Private Research Institute Clinical Economics (Ringgold iD: 587744), Schwarzenbergstr. 135, 89081 Ulm / Germany
E-Mail: mindset@clinical.economics.eu

Introduction

We discussed the congruence of our initial healthcare goals of with the observed outcomes in our cancer patients about 30 years ago [1][2][3][4]. Many of our expectations were based on a maximum achievable, but often unrealistic, goal. Instead of selecting a safe strategy enabling an optimal quality of life, our focus was directed towards the prolongation of survival confirmed in experimental, randomized controlled trials (RCTs). It took years of strategic development to reach a consensus that fulfilled patient expectations. In all our considerations to prove the suitability for everyday application, it was important not to exclude effects that result from the relationship between doctor and patient, even if we still did not have a method to describe these effects [4][5][6]. Today we are facing a similar challenge when we make decisions without having previously checked different patients' chances of achieving the desired goals.

Here we present three examples confirming that the desired goals of healthcare are frequently not achieved. Therefore, it makes sense to establish a system that predicts the probability of success of planned strategies. The front-end processor (FEP) is a system developed by engineers to confirm the congruence of manufactured steel with its constructors' goals [7]. We are convinced that this system can be applied to healthcare provision to confirm that the desired results can actually be achieved by the chosen strategy.

Methods

Strategy

Three valuable strategic clues facilitated the methodical approach. Through our cooperation with former members of the *Hochschule für Gestaltung* (Ulm Academy of Design, 1953-1968), we learned that some solutions may benefit from theoretical considerations. One example of theoretical considerations is the "Form Follows Function" (FFF) rule of American architects and designers [8]. This rule states that the chosen form, i.e., the structure of a product or concept, must correspond with the expected function of that product or concept. If this correspondence is missing, the suitability of the product or concept may require further examination. Other examples of theoretical considerations are part of the valuable legacy of Albert Einstein (*1879 in Ulm), "A problem cannot be solved with the same mindset that caused the problem.". A third theory was proposed by Sir Archibald Cochrane, who described three questions (diagnostic or

therapeutic) to be answered before introducing a new intervention in everyday healthcare, “Can it work? Does it work? Is it worth it?” [9]. In accordance with this proposal, we developed the three-dimensional strategy shown in Table I.

Cochrane questions	Can it work?	Does it work?	Is it worth it?
Outcome dimensions	Efficacy or Proof of Principle (objective PoP)	Real-World Effectiveness (objective RWE)	Value individual or societal (subjective)
Study conditions	Experimental study condition (ESC)	Observational, RWC with systematic evaluation of data	Observational, RWC, no systematic but individual evaluation of data
Research perspectives	Clinical Research	Health services research	Economic Research
Functions	Demonstration of Proof of Principle	Confirmation of Real-Word Effectiveness	Comparison of costs and consequences
Forms (structures)	Explanatory or interventional study	Pragmatic or observational study	Complete economic analysis
Tool	Randomised Controlled Trial (RCT)	Pragmatic Controlled Trial (PCT)	Cost Effectiveness Analysis (social and individual)

Table I. Three-dimensional strategy for description of Proof of Principle, Real-World Effectiveness, and Value. A possible answer to the three questions of Sir Archibald Cochrane. ESC: Experimental Study Condition. PCT: Pragmatic Controlled Trial. PoP: Proof of Principle. RCT: Randomized Controlled Trial. RWC: Non-experimental Real-World Condition. RWE: Real-World Effectiveness. References [10][11][12][13][14].

This Cochrane strategy includes the description of three different outcome dimensions, the proof of principle (PoP), the real-world effectiveness (RWE) and the value. The three outcome dimensions require three different study conditions, one experimental study condition and two observational study conditions either with or without the systematic evaluation of data. These three study conditions reflect the expected results from three different perspectives, those of clinical research, health services research and research in economics. In summary, the different dimensions, i.e., different study conditions, different outcome dimensions and different perspectives, are characterized by different functions, i.e. the demonstration of proof of principle, the confirmation of real-world effectiveness and comparison of costs and consequences and can be described with different tools, i.e. RCTs, PCTs, and different forms of cost-effectiveness analyses (CEAs).

RCTs describe the efficacy, the PoP under experimental study conditions. Various tools are available to describe the subjective value. An appropriate tool, the PCT for assessment of the RWE had to first be developed before Cochrane’s theory could be realized. Therefore, a short description and instruction for completion of a PCT is included in the second part of our results. The complexity of the three-dimensional outcome assessments emphasizes the importance of the exact study question. We recommend the following four steps from the study question to the correct answer.

Tactics

Application of the FEP requires comparison of two outcomes. These are: a theoretical expectation or the ‘front’ result

generated under experimental study conditions, i.e. the demonstration of a PoP in an RCT. In contrast, the ‘end’ result, i.e. the demonstration of real-world effectiveness (RWE), needs to be generated under the conditions of daily healthcare provision. Otherwise, the theoretical or experimental expectation cannot be confirmed by clinically relevant observations. To demonstrate the practical application of the front-end processor (FEP), we present three examples:

(1) *The first example describes the path from the simple study question (front information) to the mathematical confirmation of the expected answer (end information).*

The quantified loss of information between from the posed*study question* and the mathematical confirmation of the hypothesis was derived from Meret Phlippen’s medical thesis, "Progress in Evidence-Based Medicine: concordance of study hypothesis, design, and statistical test" [15]. She cooperated with six international teams that analyzed predefined validity criteria in 119 publications (see contributing international team). The four-step strategy – from the study question stated in simple language to the mathematical confirmation of the scientific hypothesis – only became recognizable after the initially independent questions in the dissertation were re-arranged in logical order from the clinical question to its mathematical confirmation.

(2a) *The second example demonstrates that three different healthcare conditions are needed to generate three different outcomes that can be detected with three different tools.*

The initial theory for the later explanation of the front-end processor was developed in a seminar together with students and professors of the Universidade Federal Fluminense (UFF) in Niteroí / RJ / Brazil. The aim was to describe the consecutive considerations that scientists make when conducting an experimental study and practitioners make when caring for patients under everyday conditions [5][6]. Within the next decade, it became clear that the practical implementation of a non-experimental study is like squaring the circle, as two conditions need to be met: care needs to be offered and delivered under unstructured conditions because otherwise real-world effects cannot be described, and second, the outcomes need to be analyzed under structured conditions to avoid errors and biases. The application of the FFF rule to describe the two traditional conditions in healthcare provision (experimental or pragmatic) demonstrated a third condition, which combines the characteristics of the two traditional conditions, the structured experimental condition and the unstructured condition of everyday care. The 14 criteria that characterize the three conditions of healthcare provision are presented in the research section.

(2b) *The development of the PCT requires the combination of two seemingly contradictory conditions.*

In theory, this implementation of the ‘squared circle’ into a reproducible study design could be achieved by exchanging the experimental randomization with the strictly structured but non-experimental Bayes’ statistics. In a PCT, each patient receives the same care that he or she would receive outside of a study. This condition is necessary to describe the effects of everyday healthcare. All additionally required data is collected through cooperation between the administrative staff of the facility and the individual information technology (IT) system of each PCT under the supervision of nurses and physicians. These tasks include the documentation of each patient's endpoint-specific risk profiles, the therapies performed, the reasons for and changes in response to a change in the healthcare strategy and the outcomes achieved.

Each piece of information is formulated as a question based on a stored algorithm to discover implausible answers through technical control mechanisms.

(3) *Quality criteria (front information) of tests and their psychological effects (end information) confirm the congruence of expectations, facts and outcomes of 2 x 2 contingency tables.*

Traditional 2 x 2 contingency tables are established in the health sciences to confirm the quality of a test by the reference method. According to the rules of evidence-based medicine, four types of tests can be differentiated. David Sackett's rule of thumb suggests that high test sensitivity can exclude (snout) and high specificity can include (spin) an investigated condition [16]. Most tests are suitable for either inclusion, i.e. confirmation, or exclusion of an examined condition, like a disease. Only a few tests are suitable to confirm both inclusion and exclusion, and, hopefully, a 'neither-nor test' will not be used.

The indicators for these quality characteristics of a test are calculated by the traditional analysis of the 2 x 2 contingency table. In this traditional evaluation, the attending physician is initially the recipient of the information. On the basis of this data, the physician selects the appropriate information, which he communicates with the patient in a second step. This second step is based on exactly the same numbers (forms / structures) as the first step. However, in the second step, he flow of the communication takes place in reverse direction. The recipient of the information is now the patient. In addition, the two functions of communication to the doctor and to the patient are absolutely different. In the first step, the objective information transmitted to the doctor will influence his concept of the further care strategy. In the second step, each patient receives objective information from the doctor. This information concerns the notification of the objectively existing risk from the doctor's perspective. But no patient is interested in the objective risk. Patients are interested in the concrete significance of this news for her/his own future health. In other words, the doctor's objective message is perceived subjectively by the patient. However, almost no patient can distinguish between objective risk and subjectively perceived prognosis. In our experience, this also applies to doctors when they become patients themselves. This strategy is also in line with the FFF rule of the designers. If the form/structure and function of a relationship are exchanged, the statements that can be derived from this relationship will also change. In simple terms, the calculation of the traditional table is rotated by 90^0 without changing the numbers of the table. Figure 1 illustrates this rotation using the example of X-ray mammography as 'test' and the histopathologic examination of a biopsy specimen as 'reference method'.

⇨ A positive LR describes the proportion of positive biopsies among all biopsies in the subgroup of patients with a positive mammogram

⇨ A negative LR describes the proportion of positive biopsies among all biopsies in the subgroup of patients with negative mammography.

⇩ A positive LR describes the proportion of positive mammograms in all mammograms in the subgroup of patients with a positive biopsy.

⇩ A negative LR describes the proportion of positive mammograms among all mammograms in the subgroup of patients with a negative biopsy.

		Result of histopathologic exam of biopsy (Reference)		
		Cases with positive biopsy result	Cases with negative biopsy result	Total number of patients
Results of imaging method (x-ray, computerized tomography, ultrasound, Positron emission tomography)	Cases with positive result of imaging method			
	Cases with negative result of imaging method			
	Total number of investigated cases			

Figure 1. Graphical analysis of a 2 x 2 contingency table using the example of breast cancer mammography and biopsie. Identical data can be analysed differently depending on the direction of the transmitted information and thus fulfill different functions. In the case of a medical test, traditional analysis describes the quality of the test to its user. The alternative analysis quantifies the health risks that can be communicated to the patent.

In the traditional calculation, the +LR compares the frequency of having a positive X-ray result confirmed or not confirmed by the examination of the tissue in mammography. The - LR describes analogously in the case of a negative mammogram its agreement with the result of the tissue examination. Accordingly, the comparison of the two results (x-ray and tissue examination) describes the quality of the test.

Results

(1) *Four steps from the study question (front) to the mathematical confirmation (end).*

The four-step strategy shown in Figure 2 includes [a] the clinical question (aim of study) in plain language, [b] the description of the compared targets, [c] the formulation of the study hypothesis and [d] the statistical test confirming the study hypothesis (Figure 2).

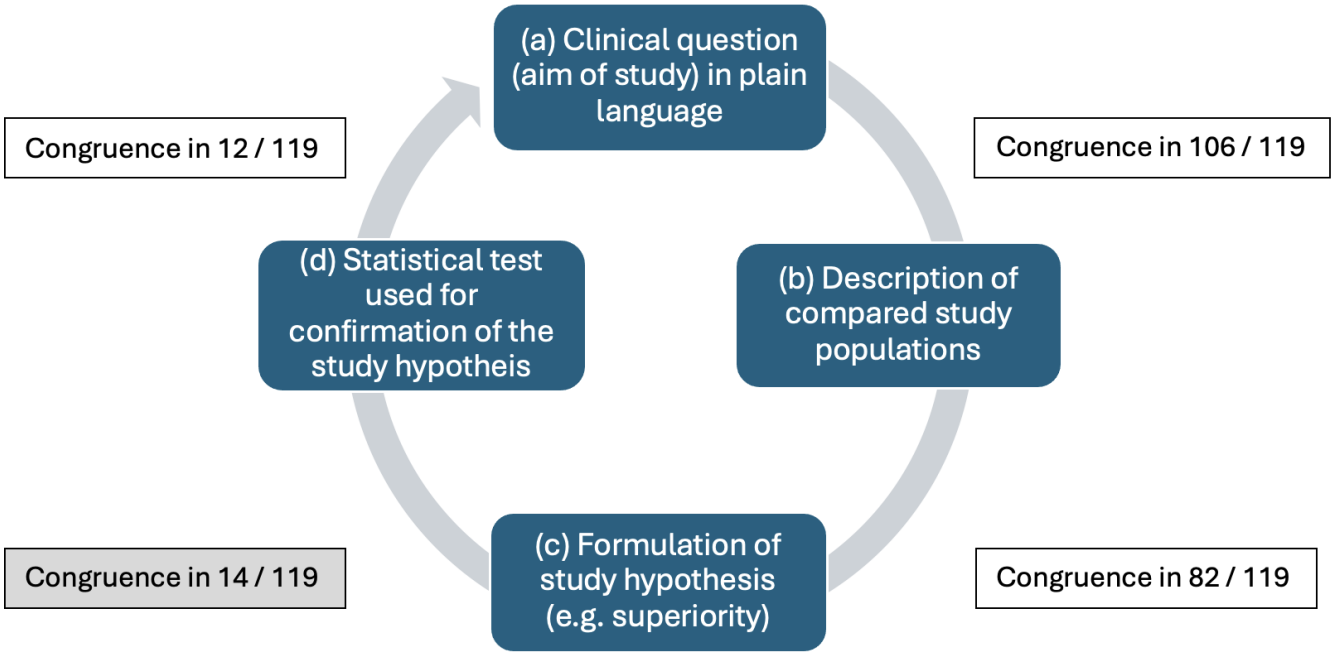

Figure 2. The way from (a) the study question in plain language to (d) the description of the applied statistical test in clinical studies. The total numbers of compared studies (n = 119) and the numbers of studies with congruent content in the steps (a and b, b and c, c and d and d and a) are shown. The least congruence (marked grey) was observed between step c and step d. Expanded reference [15]

In each multi-stage process, part of the transmitted energy or information will be lost in the transition from one stage to the next. In the clinical-question strategy described, we found differences in the transition from one of the four stages to the next in 11%, 20%, 57% and 2% of the question. When applying stringent criteria, the initial question was congruent with the mathematical confirmation in only 10% of the 119 studies.

(2a) *Fourteen features describe the differences in the 'front' condition of care (PoP) and the end conditions of care (RWE) of the front-endpoint processor.*

Here we present the two functional and twelve formal (structural) items that characterize both the structured experimental study condition (ESC) and the non-experimental, real-world condition (RWC). The ESC enables the description of the PoP, i.e. the expected (not the observed) outcomes of 'care as usual' (CAU). The expected outcome represents the 'front message' of information gain.

The scientific description of CAU can only succeed if the functions and the forms (structures) of CAU are specified. This specification can only succeed if the traditional way of thinking is abandoned to gain knowledge in healthcare. The RCT was a crucial milestone in the detection of outcomes in experimental studies. The effects of CAU cannot be observed under experimental conditions. CAU initially appears to occur in a state of 'natural chaos' [17]. The aim of evidence-based medicine "…is the conscious, explicit and judicious use of current best evidence in making decisions in the care of individual patients." [18].

Table II describes a proposal to distinguish the three conditions of care based on functional and structural characteristics.

Conditions Functions and Forms	Experimental study. Provision and analysis of care in structured conditions of an RCT*	Pragmatic study. Provision of care in un - strucured conditions but analysis in well-structured conditions of a PCT*	Unstructured care in day-to-day conditions with evaluation of outcomes in individual patients
Functions			
Provided type of care	Experimenal care	Pragmatic care	Pragmatic care
Gained information	Proof of Principle (PoP)	Real-World Effectiveness (RWE)	No systematic evaluation
Forms / structures			
(c) Impact of patient risks on results	Random distribution of risks	Endpoint-spec. risk classification	No formal control for risks
(d) Recording of applied interventions	According to study protocol	Classified by type and intensity	No standardized recordings
(e) Study protocol	Required	Required	None
(f) Approval by IRB	Required	Required	None
(g) Consent collection of personal data	Required	Required	None**
(h) Defined inclusion criteria	Study protocol	Study protocol	None**
(i) Defined exclusion criteria	Study protocol	None	None
(j) Patient health problems	Single and specified	Multiple problems	Multiple problems
(k) Intervention selected by	Study protocol	Doctor/patient	Doctor/patient
(l) Research liability insurance	Required	None	None
(m) Docs experience with clinical trials	Required	None	None
(n) Consent with exp. intervention	Required	None	None

Table II. Two functions (lanes a and b) and twelve forms (lanes c through n) define three different trials and three different conditions of care: Experimental care used in Randomized Controlled Trials (RCTs) and Pragmatic care used either in Pragmatic Controlled Trials (PCTs) or in the unstructured conditions of day-to-day CAU without systematic evaluation of outcomes. A yellow background marks forms and functions that are different in the three conditions of care. Formal criteria that are identical in two of the three conditions of care are marked by a blue background. IRB: Institutional Review Board. *RCT: Randomized Controlled Trial. *PCT: Pragmatic Controlled Trial. **None: Conditions defined in the doctor-patient contract. Previous version published as preprint in reference [13][14]

When care is provided under the **CAU** conditions, the natural chaos prevailing in a RWC, but the analysis of outcomes under structured conditions, it will be possible to describe a RWE that represents the final message of information gain (Table II). The definition of CAU is addressed in more than 1000 publications. In six systematic reviews that included the term CAU in their titles, various topic-related solutions were proposed without a valid definition of CAU [19][20][21][22][23][24]. Although the results in Table II describe the forms and functions rather than the definition of CAU, it indicates comprehensible differences among three different care conditions. The initial differences were observed when we compared the steps completed in experimental studies or in CAU [5][6] and were further supplemented by information we gained through discussions concerning the terminology conflict [12].

(2b) *The step-by-step description of combining the two seemingly contradictory conditions*

The solution of contradictory conditions can succeed if it can be confirmed that a contradiction was only assumed but does not really exist. This constellation applies to the design of the PCT, which can be used to prove the suitability of healthcare services for everyday use.

The quality of any clinical study will be poor without a precise definition of the inclusion and exclusion criteria[5]. It will be influenced by the investigated clinical problem, the selection of the inclusion criteria, and the selected study endpoints.

Inclusion criteria are identical in corresponding experimental and pragmatic trials, but study endpoints differ in RCTs and PCTs. Experimental trials can assess the results of the primary endpoint in an experiment, while an observational method, such as a PCT, can assess the results of several endpoints (e.g., main health problem, major side effects, monetary costs) as a result of CAU.

The patients need to be stratified according to their individual risk profiles related to each of the assessed endpoints. This stratification requires an algorithm that defines different risk classes. A similar algorithm is needed to define either single treatments or classified types of treatments.

All interventions not included in one of the defined intervention groups will be summarized in a mixed control group of different interventions. This mixed group may function as a reference group to compare the superiority of each of the specified target interventions (Figure 3).

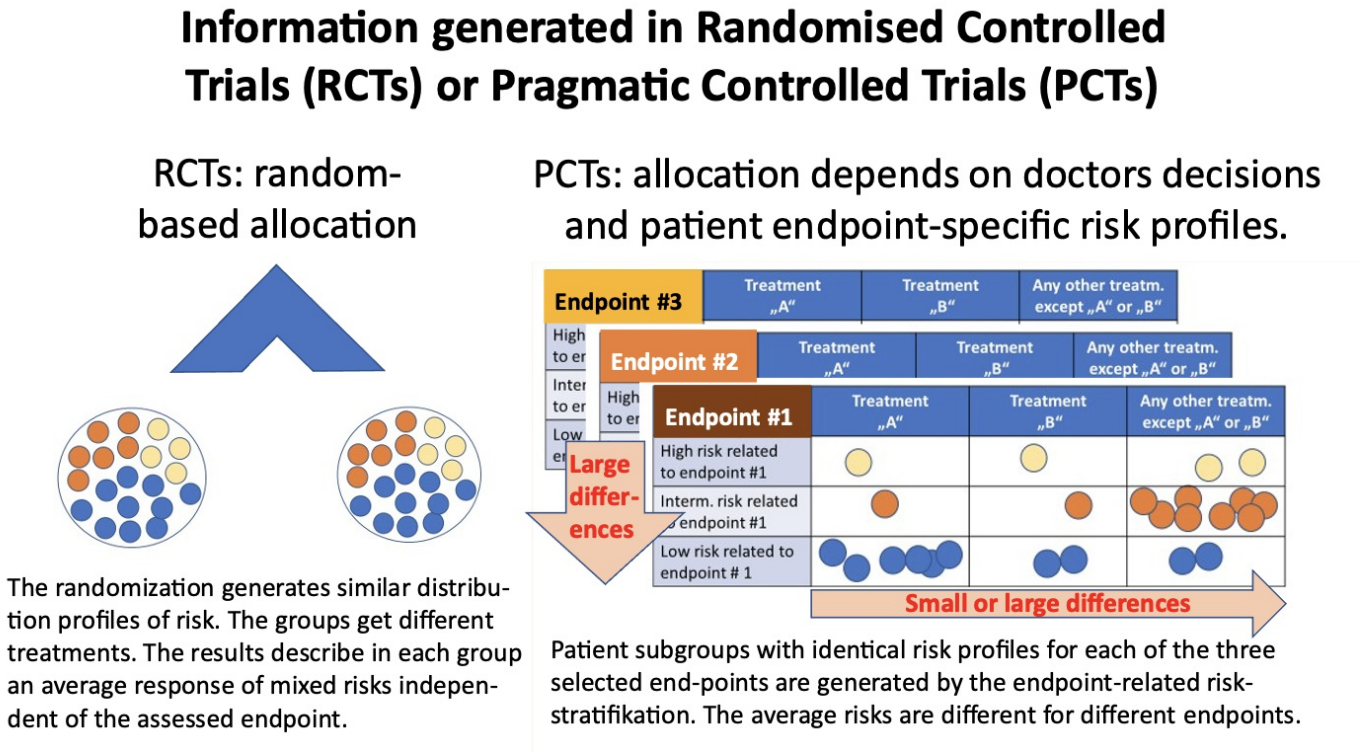

Figure 3. Forms and functions of experimental RCTs and pragmatic PCTs. Modified from [10] and published in [13][14].

The primary analysis includes the horizontal comparison of results in patients within the same class of risk profiles. The vertical analysis of the same data can detect additional information deriving from differences in risk profiles. In summary, the amount of information obtained in a PCT (efficacy or PoP) is probably higher than the amount of information obtained in a RCT (effectiveness) because RCTs include more unconfirmed assumptions about the composition of the sample studied (number, types and interdependencies of risk factors) than PCTs.

A frequent methodological problem results when two different endpoints (e.g. survival and quality of life) are reported in the same study. Most oncological studies use the design of a RCT to assess treatment effects on an objective endpoint,

such as mortality, and a subjective endpoint, such as health-related quality of life (hrQoL) [25]. Both effects were assessed under experimental study conditions in a RCT and describe efficacy data [12]. The unidimensional effect of treatment on mortality can be reliably assessed in a simple experimental trial. The expected differences in survival will be smaller than the expected effects on specific QoL dimensions, such as pain or mobility. The observed effects on these selected QoL endpoints can frequently be missed when assessed under ESCs instead of RWCs [12]. Assessing both endpoints, survival and quality of life, under everyday conditions, i.e. in a PCT, offers the simplest solution to the problem. The implementation of this PCT requires the following steps:

1. A proposed solution should be suitable for all patients who
    1. suffer from a common current health problem,
    2. are exposed to other risks in addition to the current problem, which we summarize as the concept of a patient's individual risk profile, and understand that
    3. every physician will base his care strategy on the best of his knowledge and belief in the information he has gained from a and b.
2. We need to define a study population characterized by
    1. The precise definition of inclusion criteria to describe the investigated population. The reproducibility of results depends on the precision of all definitions in this trial because all patients who meet the inclusion criteria must be cared for.
    2. Exclusion criteria do not exist in PCTs; all members of this population need to be included in RWCs.
    3. Data protection regulations need to be strictly followed. Written informed consent is necessary due to the systemic data collection.
3. Definition of the three study endpoints (for reasons of feasibility) include
    1. main outcomes to be assessed,
    2. main side effects to be recorded and
    3. the cost of healthcare should only be relevant for the decision if the objectively demonstrable effects confirm different equivalent strategies.
4. Definition of the endpoint-specific risk profiles (ESRP) of individual patients
   This can be achieved by national scientific associations to guarantee the comparability of risk profiles within and among different cultures.
    1. Prepare a separate list of risk factors that influence one of the assessed endpoints, i.e. the list of endpoint-specific risk factors.
    2. Design the algorithm that describes the criteria for classifying endpoint-specific risk profiles that differentiate high-risk from intermediate-risk and low-risk patients. The IT system in your PCT will use this algorithm for the allocation of individual patients to the appropriate risk groups separately for all assessed endpoints and the final evaluation of

study results.

3. Prepare a common list of all risk factors.
4. Use the common list to assess the individual risk profile of each patient before study entry. These individual risk profiles will be stored in the study IT system.

5. Selection and use of the best possible treatment

    1. The physician selects the best possible treatment based on the health problem, the risk profile and the preferences of his individual patient.
    2. All treatment components (e.g., doses, time schedule, initial treatment) need to be described according to the individual study protocol.
    3. Any change in the initial strategy needs to be reported (according to the individual protocol) together with the date of change.

6. Stratify patients to the target groups or the control group.

    1. To manage the huge variety of best possible treatment strategies, one or two treatment strategies of interest need to be defined.
    2. Patients who received a treatment of interest will be stratified to one of these specified target groups (two target groups can be optimal to avoid too many subgroups).
    3. All other patients will be stratified to the risk-specific mixed control groups.

7. Intervals for assessment of early and late outcomes

    1. Two intervals between a patient's study entry and assessment of the early and late responses are recommended (e.g., 2 weeks and 2 months after study entry).
    2. Remind the attending practitioner of the requested feedback.

8. Evaluation of results

    1. The **PCT** results are evaluated and reported according to the triple-stratification (endpoint, endpoint-specific risk profile and intervention) strategy. Further details are discussed in expert rounds.

(3) *Quality criteria (front information) of tests and their psychological effects (end information) can be calculated from the same contingency table.*

We have recently shown that the traditional 2 x 2 contingency table contains additional information that has not been described so far.

Table III a: Conventional 2 x 2 table

		Reference method (Sufficient condition)		
		Biopsy conf. bc.	Biopsy not conf.	Total
Test method (Necessary condition)	x-ray pos.	7	449	456
	x-ray neg.	3	4541	4544
	Total	10	4990	5000

+ LR = (7 / 10) : (449 / 4990) = 7.78

- LR = (3 / 10) : (4541 / 4990) = 0.33

Table III b: 2 x 2 table with exchanged coordinates

		Test method (necessary condition)		
		x-ray pos.	x-ray neg.	Total
Reference method (sufficient condition)	Biopsy conf. bc.	7	3	10
	Biopsy not conf.	449	4541	4990
	Total	456	4544	5000

+ LR = (7 / 456) : (3 / 4544) = 23.25

- LR = (449 / 456) : (4541 / 4544) = 0.99

Table III. Different types of information can be derived from the analysis of traditional 2 x 2 contin-gency table before and after exchange of the coordinates. Both tables show data of the Breast Cancer Surveillance Consortium tool (https://tools.bcsc-scc.org/BC5yearRisk/calculator.). The data describe the numbers of women with positive or negative X-ray mammography results and positive or negative histopathological assessments tissue samples collected by fine-needle biopsy. Modified from [14]

Part (III a) shows the traditional arrangement of the table. A positive Likelihood Ratio (+LR = 7.78) means that a positive x-ray mammography (test) will be observed 7.78-fold more frequently in women with a positive biopsy (confirming breast cancer) than in women with a negative biopsy (not confirming breast cancer). A negative Likelihood Ratio (- LR = 0.33) means that a negative X-ray mammography (test) will be observed 0.33-fold less frequently, i.e. in 1/3 of women with positive biopsy (i.e. in 3 out of 10 women; 33%) than in women with a negative biopsy (i.e. 4556 out of 4990 women; 91%).

Part (III b) shows the same table after the exchange of the test method (necessary condition) and the reference methods (sufficient condition). A positive Likelihood Ratio calculated from the inverted table (+LR = 23.25) means that the communication of a verified breast cancer to the patient will be observed 23-fold more frequently in women with a positive x-ray mammography (suggesting breast cancer) than in women with a negative X-ray mammography (no sign of breast cancer). A negative Likelihood Ratio (- LR = 0.99) means that a negative biopsy (i.e. a biopsy not detecting cancer) will be observed with the same prevalence in women with negative or positive X-ray mammography.

In summary, the + LR calculated from the *conventional table* is strong enough (marked yellow) to confirm the investigated condition, i.e. breast cancer. This test can be used to confirm the investigated disease. In contrast, the – LR is too weak (i.e. similar ratios in biopsy positive and biopsy negative samples) to exclude the investigated disease. That means that this test cannot be used to exclude the investigated disease. In the table with exchanged coordinates (Table III b), the + LR is strong enough to induce anxiety and fear as the communication of a positive X-ray mammography will be 23-fold more often be verified by the biopsy in the x-ray positive group than in the x-ray negative group. In contrast, the – LR is too close to the indifference value (- LR = 1) that describes identical prevalences of biopsy results X-ray mammography positive and negative groups of women.

This new information can be quantified by exchanging the positions of the two coordinates in a traditional 2 x 2 contingency table (Table III a/b) for calculation of the positive (+LR) and negative (-LR) likelihood ratios. The two conditions are the test method or necessary condition and the reference method or sufficient condition for calculation.

In the analysis of a conventional 2 x 2 contingency table, the physician or scientist derives information from the test and the reference method, both of which were performed on the patient. In this case, there is a flow of information from the patient to the physician or scientist. When exchanging the necessary condition (i.e. the test method) and the sufficient condition (i.e. the reference method), the generated information presented in the 2 x 2 contingency table flows from the physician to the patient, i.e. in the opposite direction compared to the flow of information calculated from the traditional contingency table.
Based on informal experience, we hypothesized that the communication of unwanted information (bad news) from the patient perspective may induce anxiety and fear, while the communication of the expected information (good news) may increase the patient's sensation of perceived safety, i.e. the subjective perception of objective risks [26][27].

The two traditional and independent properties of a test, the confirmation or exclusion of an investigated condition, are shown in Table IIIa. We recently reported that the psychological effects induced by informing the subject of his/her test results can be quantified without collecting new data. New information can be obtained simply by exchanging the two test conditions, the necessary condition (or test method) and the sufficient condition (or reference method) [26]. Using mammography as an example, the traditional way of obtaining information corresponds to a flow of information from the patient to the doctor. The doctor gains new information, i.e. the diagnosis from data collected from the patient. The new way to obtain information corresponds to a flow of information in the opposite direction i.e., from the doctor communicating the test results to the patient. This news can be desired (good news) or feared (bad news) from the patient's perspective. Using data from ten real and assumed test conditions, we were able to show that the risk of confirming an undesirable test result by the reference method is significantly higher in the test-positive group (up to 5900 times) than in the test-negative group. In the example (Table III b), the risk of a confirmed breast-cancer diagnosis is about 23 times higher in the test-positive group than in the test-negative group.

Discussion

This thesis aims to make a methodological contribution to the proof of the quality of scientific products. In previous papers, we identified and solved the 'efficacy / effectiveness' terminology conflict by applying the FFF-rule (Form Follows Function) of American designers and architects [12][13]. We want to show that another rule, the front-end processor (FEP) developed for quality controls of industrial products, is also a useful tool for assessing the quality of scientific products [7].

The designers' FFF-rule was used to confirm the congruence of the study question and the scientifically confirmed answers in clinical trials. This confirmation was shown using three examples: the agreement of the clinical study question with the mathematical confirmation of the study hypothesis; the agreement of the criteria, which are identical to or different from proof of principle and real-world effectiveness; and, finally, the agreement of objective risks with their subjective risk

perception influenced by unavoidable information effects.

The expected outcomes can hardly be confirmed if the definition of the predicted and the confirmed outcomes are described using different criteria [28]. Such problems can be detected in several papers. One of the causes of these asymmetrical effects is the disregarded variance of the subjective perceptions of objective risks [26][29]. It is the joint task of citizens and their political representatives to articulate justified patient demands. We have dispensed with the nomenclature of "necessary" and "sufficient" conditions used in mathematics here, because in health care the INUS condition, i.e., "an insufficient but necessary part of a necessary but insufficient condition" must also be taken into account [30][31]. The prospective quantity and quality of goals is indispensable if the financing of a healthcare system is not based on the services provided, but on the added health value achieved. Numerous publications describe contributions to results-oriented financing of healthcare services. However, we could not find any article in which the three methodological prerequisites, the risk profiles of the compared patients, representative samples of the CAU and the concrete objectives of healthcare were described. Without knowledge of these three factors, the resulting subjective perceptions of objective risks cannot be assessed [32]. Accurate scientific questions require the consideration of three risks: the imprecise study question, the inappropriate interpretation of reliable results and the distorted perception of objective risks through anxiety- or fear-inducing information. Our presentation describes the importance of a precisely stated study question because the answer to it can be right or wrong, but answers to imprecise study questions are almost always wrong. Therefore, the effort expended to formulate a precise question is the lesser evil.

Three forms of bias (A, B, C) influence the decisions of three professional groups (a, b, c) designing and delivering healthcare. These three forms of bias are (A) the incorrectly formulated study question, (B) the wrong interpretation of correctly assessed outcomes and (C) the effects of bad news on decisions made by (a) scientists, (b) practitioners and (c) policy makers.

In the first part we discuss the significance of the multiple losses of information that occur on the way from the clinical question to the mathematical confirmation of the hypothesis. This demonstrates the need for correction. The loss of information from the first to the fourth transition is 11%, 20%, 57% and 2%. These figures confirm that a significant part, about 90%, of the information can be lost on the way from the question to the mathematical confirmation of the study hypothesis. When interpreting these data, two effects should be considered. With any multi-step process, loss occurs at every transfer. In our case, part of the information is lost. When rearranging the sequence of the questions that were asked in the dissertation project, we noticed that different names were sometimes used to describe the same target. This resulted in minor quantitative deviations, which did not influence the interpretation of the results. Most of the information loss (57%) was caused for security reasons and is easy to correct. In clinical epidemiology, there is a consensus on the use of one-sided or two-sided statistical tests [33][34]. A two-sided statistical test has two functions. It detects an existing difference and indicates the better solution. One-sided tests are sufficient to confirm the difference if the better solution is already known (e.g. in a placebo-controlled study). They should not be misused to enhance unsuccessful experiments.

In the second part, we explain that almost every patient has an individual health-risk profile because most patients suffer from more than a single health problem. Randomization can ensure the equal distribution of single risks under ideal study

conditions but cannot compare the outcomes of healthcare under the conditions of the ‘natural chaos’ prevailing in everyday healthcare. This natural chaos cannot be avoided

This problem can be solved by an algorithm that defines the risk factors which classify all study participants as either high-risk, intermediate-risk or low-risk subjects. Second, the risk profile of each individual patient depends on the assessed endpoint. The risk profile related to myocardial complications is different from the risk profiles of frequent side effects, like allergic reactions. The algorism defining risk classes needs to be defined separately for all PCT endpoints. This problem can hardly be solved by multiple randomization, but by the Bayesian principle, which appropriately describes the effects of different risks on defined endpoints.

In the third part, we show that the communication of unwanted messages in the media can induce considerable uncertainty and fear, which have a significant influence on the ability to make rational decisions. The rate of unconfirmed and unsolicited messages is probably high, especially in international conflict situations. Most of the population is exposed to this information. If scientists and physicians want to maintain self-critical control of the own professions, they need to remember that many results and diagnoses are perceived as prognoses, although hardly any measured result and no diagnosis justifies a prognostic statement without additional specification or confirmation. If a patient's clinical findings do not match the test results, not the patient, but the test should be subjected to a quality check. Perhaps we should more critically consider the power of information [35]. We need trustworthy institutions capable of verifying the validity of undesirable information disseminated by the media. If the validity of communicated information cannot be confirmed, undesirable effects could be reduced by appropriate comments from such institutions.

In summary, we assume that conscientious doctors make their decisions based on their patients’ perceived risk profiles. The method for optimizing healthcare is not new, but the understanding of natural processes has increased. Today, in many situations – except for many psychological effects inside and outside of evidence-based medicine – we can explicitly define the risks that could previously only be perceived implicitly. We can also add two comments to Sir JA Muir Gray's statement, "Scientists make decisions – policy makers take decisions" [36]. Scientist’s decisions should be based on objective RWE data whenever available, while policy makers’ or mangers’ decisions will always be based on subjective values due to the multiple aspects that need to be considered. Scientists should first submit their proposed decisions – sometimes from different perspectives with different conclusions – followed by the policy makers’ or managers’ selection of the final decision.

The importance of these methodological considerations only becomes clear when one realizes their consequences. One of the important lessons that we can draw from our reflections concerns the lack of scientific precision of some health policy definitions. These entail cost-intensive measures without having compared the ethical, medical, epidemiological, economic, legal, and political risks with the added value achieved. Front-end processor (FEP) violations influence almost all balancing decisions in healthcare. A constructive discussion of unexplained observations such as the lack of international consensus on the recommendations in clinical guidelines [37][38] and an update of the concept of clinical economics (CLINECS) [39][40][41][42] will result in a higher economic added value than any turnover-based strategy. Prudent economists use compatible stabilizers from neighboring disciplines to increase their own chances of success.

Notes

Contributors. Each team analyzed the validity of 20 publications from their journal of choice:

- *New England Journal of Medicine (NEJM)*: Karthik Ghosh M.D., internal medicine and obstetrics / gynecology and Amit Ghosh M.D., internal medicine/nephrology, Mayo Clinic, Rochester/MN, USA
- *Malaria Journal (MJ)*: G. Oscar Kamga Wambo M.D., Msc., internal medicine, public health, social medicine, infectious diseases, private practice, internal medicine, Berlin/Germany.
- *Clinics:* Tania Gouvêa Thomaz, MD., and Cristiane Moraes PhD, Medical Faculty Universidade Federal Fluminense (UFF) Niterói/RJ/Brazil.
- *Annals of Surgery:* Valerio Balassone M.D., digestive surgery, Ospedale Pediatrico Bambino Gesu, Rome/Italy.
- *Pediatrics:* Paola Rosati M.D. MSc, pediatrics/epidemiology, Ospedale Pediatrico Bambino Gesu, Rome/Italy.
- *Journal of the American Medical Association (JAMA)*: MSP and FP (authors of this paper).

Statements and Declarations

This project did not receive any financial support.

The authors have no conflicts of interest to declare.

References

1. ^*Porzsolt F, Tannock I. Goals of palliative cancer therapy. J. Clin. Oncol. 1993;11:378-381.*
2. ^*Sigle JM, Porzsolt, F. Practical aspects of quality-of-life measurement: Design and feasibil-ity study of the Quality-of-Life-Recorder and the standardized measurement of quality of life in an outpatient clinic. Cancer Treatment Reviews 1996;22:75-90.*
3. ^*Porzsolt F, Wirth A, Mayer-Steinacker R, Schulte M, Negri G, Suhr P, Link KH, Gaus W, Röt-tinger EM. Quality assurance by specification and achievement of goals in palliative can-cer treatment. Cancer Treatment Reviews 1996;22:41-50.*
4. a, b*Coates A, Porzsolt F, Osoba D for the Internationale Gruppe Lebensqualität in der Onkologie (IGLOO). Quality of life in oncology practice: prognostic value of EORTC QLQ-C30 scores in patients with advanced malignancy. Eur. J. Cancer: 1997;33:1025-1030.*
5. a, b, c, d*Porzsolt F, Eisemann M, Habs M, Wyer P. Form Follows Function: Pragmatic Controlled Trials (PCTs) have to answer different questions and require different designs than Ran-domized Controlled Trials (RCTs). J Publ Health 2013;21:307-313. DOI 10.1007/s10389. https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3655212/*
6. a, b, c*Porzsolt F, Rocha NG, Toledo-Arruda AC, Thomaz TG, Moraes C, Bessa-Guerra TR, Leão M, Migowski A, Araujo da Silva AR, & Weiss C. Efficacy and effectiveness trials have different goals, use different tools, and generate different messages. Pragmatic and Observational Research 2015;6,47-54. Erratum in: Pragmatic and Observational*

Research, 7, 1. PMID: 27774032.

7. [a, b]*Suresh BR, Fundakowski RA, Levitt TS, Overland JE. A real-time automated visual inspec-tion system for hot steel slabs. IEEE Trans Pattern Anal Mach Intell. 1983;5:563-572. doi: 10.1109/tpami.1983.4767445.*
8. [^]*Sullivan, LH. The tall office building artistically considered. Lippincott's Magazine 1896;57, 403-409. Reprinted in Inland Architect and News Record 27 (May 1896), pp. 32-34; Western Architect 31 (January 1922), pp. 3-11; published as "Form and Function Artistically Con-sidered" The Craftsman 8 (July 1905), pp. 453-58. https://www.giarts.org/article/tall-office-building-artistically-considered-excerpt. Last access oct 22, 2019*
9. [^]*Haynes B. Can it work? Does it work? Is it worth it? The testing of healthcare interventions is evolving. BMJ. 1999 Sep 11;319(7211):652-3. doi: 10.1136/bmj.319.7211.652. PMID: 10480802; PMCID: PMC1116525.*
10. [a, b]*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.*
11. [^]*Porzsolt F, Weiss Ch, Weiss M, Jauch K-W, Kaplan RM. Demystification – a Solution for As-sessment of Real-World Effectiveness? Trends Med 2020;20:1-2. doi:10.15761/TiM.1000231*
12. [a, b, c, d, e]*Porzsolt F, Wiedemann F, Phlippen MS, Weiss C, Weiss M, Schmaling K, & Kaplan RM. The terminology conflict on efficacy and effectiveness in healthcare. Journal of Comparative Effectiveness Research 2020;9:1171-1178. PMID: 33314965. https://doi.org/10.2217/cer-2020-0149*
13. [a, b, c, d]*Porzsolt F, Weiss M, Weiss Ch. Applying the Rule of Designers and Architects "Form Fol-lows Function (FFF)" Can Reduce Misinterpretations and Methodical Shortcomings in Healthcare. Trends Gen Med. 2024; 2(1): 1-7. ISSN: 2996-3893*
14. [a, b, c, d]*Porzsolt F, Weiss Ch, Weiss M. An Evidence-Based Hypothesis: Doctors Do Not Make De-cisions Randomly but Based on Individual Patient's Risk Profiles. QEIOS 2024-05-21, Pre-print. DOI: 10.32388/N02184*
15. [a, b]*Phlippen MS. Progress in Evidence-Based Medicine by critical appraisal of Evidence-Based Medicine: Concordance of study hypothesis, design, and statistical test. Doctoral Thesis, Medical Faculty, University of Ulm, 2019.*
16. [^]*Sackett DL, Richardsson WS, Rosenberg W, Haynes RB. Evidence-based Medicine. How to practice and teach EBM. Churchill Livingstone, New York, Edinburgh, London, Madrid, Melbourne, San Francisco, Tokyo. First edition 1997. ISBN 0-443-05686-2*
17. [^]*Porzsolt F, Weiss Ch, Weiss M. Covid-19: Twin-Methode zum Nachweis der Real-World Effectiveness unter Alltagsbedingungen [Covid-19: Twin method for demonstration of real-world effectiveness (RWE) under the conditions of day-to-day care]. Das Gesund-heitswesen 2022;84:1-4. DOI: 10.1055/a-1819-6237. ISSN 0941-3790*
18. [^]*Sackett DL, Rosenberg WMC, Gray JAM, Haynes RB, Richardson WS. Evidence based medicine: what is is and what it isn't. BMJ 1996;312:71-72*
19. [^]*Jayedi A, Zeraattalab-Motlagh S, Shahinfar H, Gregg EW, Shab-Bidar S. Effect of calorie re-striction in comparison to usual diet or usual care on remission of type 2 diabetes: a sys-tematic review and meta-analysis of randomized controlled trials. Am J Clin Nutr. 2023;117:870-882. doi: 10.1016/j.ajcnut.2023.03.018.*

20. ^*Mirghani H, Altedlawi Albalawi I. Metabolic surgery versus usual care effects on diabetes remission: a systematic review and meta-analysis. Diabetol Metab Syndr. 2023;15:31. doi: 10.1186/s13098-023-01001-4.*
21. ^*Pascoe SC, Spoonemore SL Jr, Young JL, Rhon DI. Proposing six criteria to improve repro-ducibility of "usual care" interventions in back pain trials: a systematic review. J Clin Epi-demiol. 2022;149:227-235. doi: 10.1016/j.jclinepi.2022.05.002.*
22. ^*Lenferink A, Brusse-Keizer M, van der Valk PD, Frith PA, Zwerink M, Monninkhof EM, van der Palen J, Effing TW. Self-management interventions including action plans for exacer-bations versus usual care in patients with chronic obstructive pulmonary disease. Cochrane Database Syst Rev. 2017;8:CD011682. doi: 10.1002/14651858.CD011682.pub2.*
23. ^*Bhattacharya D, Aldus CF, Barton G, Bond CM, Boonyaprapa S, Charles IS, Fleetcroft R, Holland R, Jerosch-Herold C, Salter C, Shepstone L, Walton C, Watson S, Wright DJ. The feasibility of determining the effectiveness and cost-effectiveness of medication organisa-tion devices compared with usual care for older people in a community setting: systemat-ic review, stakeholder focus groups and feasibility randomised controlled trial. Health Technol Assess. 2016;20:1-250. doi: 10.3310/hta20500.*
24. ^*Bürge E, Monnin D, Berchtold A, Allet L. Cost-Effectiveness of Physical Therapy Only and of Usual Care for Various Health Conditions: Systematic Review. Phys Ther. 2016;96:774-786. doi: 10.2522/ptj.20140333.*
25. ^*Wiedemann F, Porzsolt F. Measuring Health-Related Quality of Life in Randomised Con-trolled Trials: Expected and Reported Results Do Not Match. Pragmat. Obs. Res. 2022;13:9-16. Doi 10.2147/POR.S350165.*
26. [a, b, c]*Porzsolt F, Pfuhl G, Kaplan RM, Eisemann M. Covid-19 pandemic lessons: Uncritical communication of test results can induce more harm than benefit and raises questions on standardized quality criteria for communication and liability. Health Psychology and Be-havioral Medicine 2021;9:818-829. DOI 10.1080/21642850.2021.1979407*
27. ^*Porzsolt F, Matosevic R, Kaplan RM. Recommendations for cancer screening would be different if we measured endpoints that are valid, reliable, specific, and important to pa-tients. Cancer Causes Control. 2020;31:705-711. doi: 10.1007/s10552-020-01309-w*
28. ^*Tullu MS. Writing the title and abstract for a research paper: Being concise, precise, and meticulous is the key. Saudi J Anaesth. 2019 Apr;13(Suppl 1):S12-S17. doi: 10.4103/sja.SJA_685_18.*
29. ^*Porzsolt F. Safety means perception of risk. J Med Safety 2016;Oct:18-24. Presented: 4th World Congress on Clinical Safety, Vienna/Austria, Sept 28-30, 2015*
30. ^*Mackie JM. Causes and Conditions. American Philosophical Quarterly. 1965;2:245–264. JSTOR:20009173.*
31. ^*Mackie JM. The Cement of the Universe: A Study of Causation. London 1974: Ox-ford University Press. Chapter 3: Causal regularities, p. 59–87.*
32. ^*Porzsolt F. Kroemer R, Silva-Sobrinho RA. The two sides of the value medal – confirmed or not confirmed by Real-Word Effectiveness? VBHC Thinker Maga-zine Christmas Edition Dec. 2022: 40-43. https://www.vbhc.nl/thinkers-magazine/vbhc-thinkers-magazine-xmas-2022/*
33. ^*Bland JM, Altman DG. One and two sided tests of significance. BMJ. 1994;309: 248. doi: 10.1136/bmj.309.6949.248.*
34. ^*Wolterbeek R. One and two sided tests of significance. Statistical hypothesis should be brought into line with clinical hypothesis. BMJ 1994; 309 doi: https://doi.org/10.1136/bmj.309.6958.873a (Published 01 October 1994) Cite this as:*

BMJ 1994;309:873

35. ^*Glasziou PP, Sanders SL. Investigating causes of heterogeneity in systematic reviews. Stat Med. 2002;21:1503-1511. doi: 10.1002/sim.1183.*
36. ^*Gray JAM. Evidence based policy making - Is about taking decisions based on evidence and the needs and values of the population. BMJ 2004;329:988–989. Editorial. doi:10.1136/bmj.329.7473.988*
37. ^*Porzsolt F, Wiedemann F, Schmaling K, Kaplan RM. The risk of imprecise termi-nology: incongruent results of clinical recommendations in clinical trials and in-congruent recommendations in clinical guidelines. 10.1136/bmjebm 2019;24(Suppl 1):A17-A18.*
38. ^*Porzsolt F, Rhoads CJ, Kaplan RM. Quantified Congruence for Cancer Treatment Recommendations from Various Countries. Int J Surg Proced 2020;3(2):136. ISSN 2517-7354*
39. ^*Porzsolt F, Kaplan RM (eds.) Optimizing Health – Improving the Value of Healthcare Delivery. Springer, New York, 2006, pp 298-303*
40. ^*Porzsolt F. Clinical Economics – It is about Values not about Money. Brazil J Med-icine Human Health 2016;4 (3). DOI: http://dx.doi.org/10.17267/2317-3386bjmhh.v4i3.1052*
41. ^*Porzsolt F. Klinische Ökonomik: Was sie ist, was sie will und weshalb wir sie brauchen? [Clinical Economics: what it is, what it wants to achieve, and why we need it]. Gesundheitsökonomie und Qualitätsmanagement 2017; 22: 131 – 134. DOI: 10.1055/s-0043-108015*
42. ^*Porzsolt F, Correia LCL. Clinical Economics is based on Real World Condi-tions not only on Ideal Study Conditions. (Editorial). Arquivos Brasileiros de Cardiologia [Brazilian Archives of Cardiology] 2017. DOI: 10.5935/abc.20170084. https://www.semanticscholar.org/reader/204a83717b93a8130116f5102e37f16814014f8c*