DECISION-MAKING IN THE NUTRITION SCIENCES: A CRITICAL ANALYSIS OF SCIENTIFIC EVIDENCE FOR ASSESSING HEALTH CLAIMS[1]

JUAN BAUTISTA BENGOETXEA[1]
[1]https://orcid.org/0000-0003-1158-1122
University of the Balearic Islands
Department of Philosophy
Palma de Mallorca, Islas Baleares
Spain
juanbautista.bengoechea@uib.cat

OLIVER TODT[2]
[2]https://orcid.org/0000-0001-9363-0543[2]
University of the Balearic Islands
Department of Philosophy
Palma de Mallorca, Islas Baleares
Spain
Oliver.todt@uib.es

Article info
CDD: 501
Received: 04.04.2021; Revised: 23.06.2021; Accepted: 25.06.2021
https://doi.org/10.1590/0100-6045.2021.V44N3.JB

[1] **Aknowledgment**

This work was supported by the European Commission European Regional Development Fund (ERDF-FEDER)/ Spanish Ministry for Science and Innovation – State Research Agency (AEI), research project FFI2017-83543-P ('Estándares de prueba y elecciones metodológicas en la fundamentación científica de las declaraciones de salud') and research project PID2020-113449GB-I00 ('Las políticas epistémicas en la ciencia reguladora: evaluación de riesgos y evaluación de beneficios').

Keywords
Health-Claim
Evidence
Nutrition
Decision-Making
RCT
Expertise

Abstract: In this paper we present an analysis of the role of randomized controlled trials (RCTs) in the regulation of health claims (claims about additional health benefits provided by foods). Currently there is a line of thought in the nutrition sciences and in regulation that data from RCTs may be able to minimize, or even make superfluous, the role played by expert knowledge in decision making. We analyze the limitations of, as well as the possible intervention of expert judgment in RCTs in pharmacology and nutrition. As a result of our analysis, we argue that both RCTs and expert knowledge are necessary for data generation in health claim regulation. We argue that as far as data generation is concerned, nutrition is more complex than pharmacology, implying that RCTs are more difficult to effectively design and execute. What the latter means is that in nutrition and health claim regulation, expert knowledge is even more important than in pharmacology.

1. Introduction: chasing scientific evidence

In this paper we propose a critical analysis of the scientific methods used to generate evidence in the *decision-oriented sciences* (hereinafter, DOSS).[2] Our motivation is the

[2] The use of scientific evidence is partially dictated by different considerations—pragmatic and epistemic, for instance—in order to make a decision on the regulation of a substance or product (like chemical compounds, novel foods, or drugs). The decision-

often cited assumption that *randomized controlled trials* (RCTs, clinical trials) on health claims are the best means—sometimes the only way—to obtain decision-relevant, objective scientific evidence,[3] meaning that the recourse to expert knowledge is in many cases considered secondary. In other words, the advocates of RCTs consider that it may be possible to make expert judgment mostly superfluous in DOSS.

Evidence serves as input for making decisions in different fields of research and policy (Achinstein 2001, Cartwright 2015). One important way of obtaining *unbiased* evidence is the *randomized controlled trial* (Hacking 1983). The objectivity produced by RCTs is based on the use of control groups (which are given a placebo or an active control), alongside the experimental ones. Trials of, for instance, drugs on the basis of the RCT design are generally considered to be able to show efficacy of the compound under study (Rubin 1974, Papineau 1994).

In observational or experimental data collection processes, a crucial condition on experiments is, especially in

---

oriented sciences are those disciplines of knowledge whose work depends upon that kind of use.

[3] We are alluding to the triple objectivity constituted by what Koskinen (2020, p. 1191), from the seven kinds of objectivity presented by Douglas (2004), summarized as *convergent*, *procedural* and *interactive* objectivity—that is, the reach of the same results via different means, the replacement of one researcher by another without altering the result, and a research community's urge to fosters diverse critical exchanges. In addition, we would like to emphasize that in this paper we focus on *enhanced function* health claims, unlike the type that, for instance, Jukola (2019, p. 4) examines in more detail: the health claims related to nutrition guidelines that might help minimize certain diseases on the level of an entire population.

social sciences and medicine, that they must be constrained in one way or another. Given that, depending on each situation, the use of causal inference methods becomes necessary, randomization has become a key concept in experimental methods. The best strategy for avoiding 'confounders', for example, is the randomization of a treatment (Illari and Russo 2014). This is a means aimed at eliminating bias due to uncontrolled differences in experimental conditions. But randomization presents problems of its own. Due to the complexity of phenomena and environmental conditions in biomedical or social contexts, randomization has been imposed as the basis for RCTs, on which most evidence-based approaches are based. Two crucial ideas of RCTs are: (1) the random application of a treatment in experimental subjects (like laboratory mice or human patients) minimizes the problem of bias and confounding, and (2) in order to establish the efficacy of a treatment, the results obtained from the test group are compared with those of the control group. In addition, RCTs can work with many individuals or with just one (trial *n*, trial *1*) and can be blind or double blind (Jukola 2019, p. 2).

The crucial issue is, however, that this idealized setup for clinical trials, which is the basis for experimentation and assessment in pharmacology (Luján and Todt 2021), is not sufficient as a basis for regulatory decision making, unless it is combined with expert judgment. This is an argument already being studied in the field of pharmaceuticals, and here we aim to analyze this same argument in the case of nutrition. Although a DOSS linked to medicine or pharmaceuticals testing can provide some equality and balance according to the standards of regulatory agencies like, for instance, the FDA (US Food and Drug Administration), we want to emphasize that several

methodological problems of various kinds have arisen (Trusswell 2001, pp. 1061-1062). We are interested here in examining these difficulties in order to shed light on the management of the notion of scientific evidence in health claim regulation (Teira and Reiss 2013, p. 209; Sackett et al. 1996, pp. 71-72).

We agree with Stegenga's argument (2015) in many respects when he examines and qualifies internal ingredients typical of clinical research procedures, and identifies three key problems in measuring the efficacy of medical interventions: use of defective or poor measuring instruments (Stegenga 2015, pp. 63-65), confusing analytical operations (interpretation of chosen parameters and of instruments used to assess whether or not an intervention modifies the parametric values) (Stegenga 2015, pp. 65-68), and the assumption that measurements within a specific experimental setup are sufficient for allowing for possible extrapolation to other fields of application (Stegenga 2015, pp. 68-70). The author claims that the response to these problems usually undermines the role of expert judgment and overestimates the effectiveness of clinical interventions.

Advocates of RCTs consider that statistical tools are impartial and impervious to particular interests that otherwise could bias the judgment of the individual expert. Then, and only after applying RCTs, the "automatic objectivity" an RCT provides should be checked against expert judgment in a continuous attempt to integrate expert judgment—the expert by her own may be misinformed, partial, biased or a 'malpractitioner'—with some kind of mechanical objectivity (Van Baalen and Boon 2015).

RCTs are critically important for health claim regulation. Health claims are statements that indicate that a food or food ingredient offers positive health effects beyond its simple nutritional value. Health claims are subject to regulation in

many countries, due to the added value which they provide. They are usually regulated by the national or a common food regulator (like, for instance, the U.S. Food and Drug Administration or the European Food Safety Authority EFSA). A primary objective of health claim regulation, apart from facilitating officially sanctioned information to consumers who want to improve their health, is to protect consumers from false or misleading claims, but also to protect innovative food products, as well as promote public health.

While there are differences between the ways health claims are regulated in, e.g., the United States, Japan, or Europe, all of those regulatory frameworks consider nutrition RCTs to provide the best evidence for regulatory authorization. That is the reason why RCTs have become crucially important for regulating many health claims. In fact, the RCT methodology was adapted from pharmaceuticals testing, and transferred to the nutrition field, in the hope of providing nutrition research with an "objective" basis. As we will see, there exists debate about the adequacy or lack thereof of RCTs in nutrition.

In order to critically analyze the generation of evidence for DOSS and particularly for health claim regulation, we have structured the text as follows: In section 2, we analyze the standard way of obtaining evidence according to the design and procedures of pharmaceutical RCTs. We examine the controversial issue of whether medical-style RCTs, given some of their methodological limitations, are in fact the best tool for generating evidence in nutrition. The high evidential status of RCTs is questionable in nutrition since data generation in this latter field turns out to be more complex than in the field of pharmaceuticals. In section 3, we focus on the obtainment of evidence in nutrition, discuss its putative objective nature, and present some crucial

differences between nutrition science and pharmacology. In section 4, we present several examples that show the limitations of RCTs for generating data in nutrition, and point to the importance of expert judgment. In section 5 we conclude arguing for the necessary complement of RCTs and expert knowledge in regulatory decision making.

2. RCTs and problems related to automatic objectivity in pharmacology

Currently there is an ongoing discussion about the epistemic foundations of RCTs (Cartwright 2010). If RCTs are conceived, as Cartwright does,[4] as tools for making causal inferences—which is not the only way to interpret them—, their impartiality can be understood as a useful byproduct for regulatory decisions. If we can objectively establish that a cause (say, a drug) is effective in obtaining an effect (say, curing a disease), the establishment of the effect can be considered independent of any interest that might be involved in the RCT (Cartwright and Hardie 2012, Hayward and Krumholz 2012, pp. 2-5, Rubin 1974, p. 700, Urbach 1985, Van Baalen and Boon 2015, Worrall 2010).

But causality presents problems. Establishing causation is a very rigid evidentiary condition. The design and execution of RCTs in order to obtain the necessary data requires a lot

[4] Making causal inferences on the basis of nutrition RCTs is fundamental for their regulation. In Europe, for instance, the regulatory agency responsible for regulating health claims (the European Food Safety Authority, EFSA) considers that RCTs are the best scientific method for generating data as basis for authorization of claims because RCT data allow for the establishment of causal relationships between ingestion of a food or ingredient and the desired positive health effect (Heaney 2008).

of resources and time. As an example, only very few health claims have been authorized in Europe due to the challenge of establishing causality between ingestion and the desired outcomes. In a perspective that closely follows Luján and Todt's methodological pluralism (2021, p. 31), we argue that in the nutrition field expert participation in decision making is *even more relevant* than in medicine and pharmacology because in nutrition many of the relevant institutions, protocols, procedures (and even stakeholders) are neither well established nor sufficiently mature. Only experts can provide the necessary knowledge and judgment that in medicine in many cases has already been institutionalized and codified.

Controlling and checking assignments such as the randomization decisions during trials, as well as the decision of what compound to submit to testing, is something that demands expert judgment, which goes beyond mere statistical calculations. Against the ideal of an "automatic objectivity", we need expertise that can manage multiple sources of evidence in order to be able to verify that causal assumptions have been fulfilled. And we also need someone who can certify that randomization has actually been successful. Without more or less subjective expert judgment, we cannot arrive at warranted generalizations from the conclusions of the trial to the target population—we cannot ascertain its external validity; i.e., the kind of validity that has to do with whether the result that is established in an RCT will be true in other, alternative implementations, as well as the real world (Jiménez-Buedo and Miller 2010; Cartwright 2010, p. 60).

In sum, the external validity of RCTs has to be verified in practice. We can see this when a regulator (the FDA, for instance) collects post-marketing reports, particularly adverse reports, from different sources, and directs

epidemiological studies that assess their relevance (Jiménez-Buedo and Miller 2010). This fact can indicate failures in the trials. In pharmacology labeling reviews and withdrawal from the market serve as checks on the external validity of FDA-approved pre-marketing trials.[5] While data generated by RCTs generally allow to establish the main effects of a drug, they cannot guarantee that all possible secondary effects are accounted for. One of the main reasons for the latter is that pharmacological trials are conducted in strictly controlled environments and with highly limited populations (the respective experimental and control groups). Real-world use of drugs (or foods, for that matter) occurs, however, in contexts that can differ significantly from the one of pre-marketing trials. RCTs are designed to try to capture relevant aspects of this real-world context, but their limited nature means that there simply cannot be any guarantee that previously unknown effects might not crop up in the marketing stage. Hence the need for post-marketing monitoring (Cartwright and Hardie 2012, p. 125).

A lot of formal and informal knowledge is acquired in the process of testing a drug, knowledge that is tested not only thru RCTs, but also by subsequent monitoring (which is more akin to an epidemiological study) after the drug obtains regulatory approval. And it turns out that expert judgment contributes to the each and every part of this entire process. Even if the decision to authorize a drug is taken on the basis of RCT data, the decision is *not an automatic one* precisely because the external validity of a trial cannot be guaranteed (Teira and Reiss 2013, p. 215).

[5] We refer to the extrapolation of results—valid for an examined population (internal validation)—from one population to an alternative population (see Illari and Russo 2014, pp. 17f).

Randomization can be considered a procedure for minimizing bias. In fact, *removal of bias* is its principal aim. Randomization prevents researchers from assigning treatments to patients according to their personal interests, so that, for instance, healthier patients get the researchers' favorite therapy. And although these unbalanced assignments may happen by chance, randomization guarantees that the assignment is not done *ex professo*. In principle, therefore, we may concede that the procedure is impartial with regard to someone's interests as well as possible biases (Teira 2013), to the extent that it is blind. That is why RCTs imply a decrease in bias.

Regulation (pharmaceutical and other) largely relies on a minimization of bias in obtaining its evidential base. Bias reduction turns out to be crucial for public policy, and it seems reasonable to adopt RCTs instead of mere expert judgment for regulatory purposes, since against those who argue that RCTs are not actually unbiased, RCTs at least provide certain minimal guarantees (Teira 2013, p. 417). As we will see in the next section, in which we analyze the case of nutrition, the same minimal (though far from complete) guarantees apply to the *monitoring* of subjects and the *extrapolation* of results (Bowen and Zwi 2005). In the nutrition sciences, it is of utmost importance both to know what the actual dietary intake of the subjects is, and to confirm that the latters have taken the adequate amount of food at appropriate time intervals and during a specified time period. Only if subjects satisfy these demands can the study be considered an adequate proof of the benefit of the food under study. Examples of a test of 'satisfaction' include blood and tissue levels of the known component or its metabolites, or the addition to the food of a marker that can be detected in blood, urine or in breath (Lähteenmäki 2013).

For our argument it is important to highlight some relevant differences between the role played by RCTs in medicine and their role in nutrition because the latter field has, from a methodological and social standpoint, a much less clearly defined scope. Institutions in medicine have been well defined and thouroughly established for a long time, but that is not necessarily so in the field of nutrition. One can easily discriminate, say, a hospital, a health system, the medical professional at its various levels, endpoints in medicine, the effects of a medication, and so on, but in many cases there are no corresponding, similarly well defined elements in the field of nutrition.

3. Nutrition: six constraints on the RCT methodology

While in principle many different types of scientific methods (like epidemiological studies) can provide scientific evidence of nutrient effects, RCTs have been the preferred source of evidence for regulatory decisions in relation to health claims. The latter can be seen, for instance, in the evidentiary hierarchies which the European regulator (the European Food Safety Authority EFSA) applies as basis for decision making; see EFSA 2011). It is however not clear if in the nutrition context RCTs always constitute the best methodology. There are features of pharmacology that do not apply in the same manner in the context of nutrition, and therefore impose constraints on the use of RCTs in the nutrition sciences. Six of the most relevant differences between drug and nutrient evaluations are (Heaney 2008, Blumberg et al. 2010):

[1] The aim of medical interventions is to cure diseases that the absence of said interventions does not

generate. In the case of nutrients, to the contrary, these prevent dysfunctions that are a consequence of a lack of intake, or of inadequate intake.

[2] It is implausible to construct clinical settings for very basic nutritional effects without generating ethical impediments to many trials.

[3] The effects of drugs are narrow and have a limited scope of action, while those of nutrients usually are polyvalent in scope and by and large act over long periods of time.

[4] The effects of drugs are often lineal—their response varies in proportion to just one variable (the dose)—while the effects of nutrients are of a *sigmoid* or curved nature.

[5] Drug effects are tested against an unexposed contrast (placebo) group; in nutrition attempting to contrast with a nutrient-zero intake group would obviously create ethical issues.

[6] Drugs are intended to act rapidly and often for only relatively short time periods; in contrast, nutrients act over very extensive periods of time in developing their beneficial effects in the human body (implying long periods of observation for measuring any positive health outcomes).

The ethical and practical reasons mentioned in the previous six points in many cases make it impossible to develop useful RCTs in the nutrition context (Goodman 2003). We would like to highlight a few points in this regard:

[A] In order to perform an RCT aiming at testing the efficacy of a nutrient, it is important to assure an adequate *contrast in intake* between intervention and control groups (Blumberg *et al.* 2010, p. 480). The intake of the control group is similar to that of placebos in drug RCTs. However,

given the low intake associated with many desirable endpoints in nutrition, any nutrient deficiency would lead to *ethical problems*, especially if it could result in serious or irreversible illness. As a consequence, and unlike observational studies—which normally assess exposures to nutrients from very low to very high ranges—, most nutrition RCTs employ a control group which typically receives an intake similar to the 'recommended daily allowance', i.e., above the threshold for deficiency states. The result is that the trials are ethically viable, but often do not test the hypothesis that really matters, namely that low intake of a particular nutrient causes negative effects (De Boer et al. 2014).

The difference between pharmacology and nutrition is critical, though. Nutrition should take into account whether inadequate intake of a nutrient might contribute to the generation of unhealthy body states. This is a relevant question because RCTs are unlikely to be able to address the role played by a nutrient with respect to diseases not yet classified in a formal medical and institutional way; that is, not registered as diseases, according to the WHO, for example.[6] This means that most of the evidence for nutrients and non-indexed or officially unrecognized diseases will continue to be derived from observational studies. In this context, the evaluation of (assumed) benefits from food is generally seen as somewhat more speculative than the evaluation of (known and well-defined) risks (from disease, for example) and, as a consequence, tends to receive lower financial support.

[6] A classic instance is the 'Manual of the international statistical classification of diseases, injuries, and causes of death (International lists of diseases and causes of death)', adopted by the World Health Organization in 1948 and systematically updated.

[B] In the DOSS context, in order to minimize risks, high *levels of certainty* are demanded in the assessment of the efficacy and safety of pharmaceuticals. In medical RCTs, the standards are indeed high. The use of those standards is justified (Tapsell 2008) (i) by the high cost of medical treatments, (ii) the risk that therapeutic decisions based on inappropriate evidence may force physicians to abandon treatment in favor of less effective therapies, and (iii) the need to balance the benefits against the risks that may accompany pharmacotherapy. The same is not true for nutrition, since nutrients are much less expensive than drugs and often exhibit a lower range of efficacy and toxicity.

[C] The previous point suggests another difference between medicine and nutrition: in pharmacology, RCTs are often conducted even if the conditions are far from ideal, simply because the health of a part of the population is at stake and there are no alternative scientific methodologies at hand to generate the relevant data. One might, for instance, decide to go ahead with a drug RCT even in the face of well known difficulties in designing the control or experimental groups. An example of the latter might be clinical trials during an emergency, like the Covid-19 pandemic, when in order to speed up development prospective vaccines might be tested on (and destined for) certain population groups only, or might be tested on fewer patients than usual.

The same is not true in nutrition. If you are unable to design an effective nutrition RCT due to, for instance, ethical or viability reasons, you still might have the possibility of recurring to data from other scientific methods (like epidemiological studies) that do not provide causal data. In the context of health claims, in which we are not concerned with food safety, this methodological pluralism (a combination of causal and non-causal methods) would also be acceptable, not only for issues related to potential health

benefits of particular foods, but also those related to the possible mitigation of population-wide risks from unhealthy food consumption habits (see Luján and Todt 2021, p. 30).

[D] The lack of certainty has led to a debate on whether the *endpoints* recognized by regulation should be taken into account or not when determining the requirements for nutrient intake (Lawrence 2013). This is a controversy that emerges from a long-standing supposition in nutrition policy, according to which we must concentrate on 'the smallest intake acceptable without developing a disease'. Its recommendations are the basis of the 'Recommended Dietary Allowance' (RDA).[7] The counterargument is that the simplistic methodological approach of 'a nutrient-a disease' is inadequate. A more holistic concept would be one that considers that (i) most nutrients act on all tissues, (ii) all tissues need all nutrients, and (iii) an inadequate intake harms all body systems (some more than others, of course) (Heaney 2008, p. 1592).

The use in nutrition studies of standards and methodologies (like RCTs) taken from medicine and pharmacological research may help to increase objectivity (in the sense of providing highly reliable statistical data on the relationship between intake and outcome). But the relevant scientific evidence never comes *exclusively* from RCTs (Neale

[7] Let us take an illustrative case studied by Heaney, "Nutrients, Endpoints, and the Problem of Proof," 1591: Vitamin D (VitD). The recommended intake up to the age of 50 is 200 international units per day. This is a sufficient amount to prevent rickets in children and osteomalacy in adults. However, it is known to be an inadequate input to maintain serum 25-hydroxyvitamin D during the winter in many North American latitudes and to produce other benefits normally associated with VitD. For RDA, see https://ods.od.nih.gov/Health_Information/Dietary_Reference_Intakes.aspx.

and Tapsell 2019). One of the principal reasons is that the establishment of causality is, as we have already seen, a very challenging requirement. In fact, there is a diversity of sources of evidence ranging from animal studies and molecular biology to the critical interpretation of observational epidemiology. It is clear that the RCT methodology might be partially inadequate to assess nutrient effects, but at the same time its ability to minimize bias is crucial for regulatory decision making.

4. Health-claims, evidence, and expert judgment for decision making in nutrition

In regulation, standards of proof are important for the taking of decisions (Steel 2015; Reiss 2015). That is because it is the standard of proof that indicate which type and level of evidence to require. In many cases standards of proof make it possible to define hierarchies among different types of evidence. Such hierarchies are a relevant tool in regulation because they indicate the type of scientific methodology to be preferred when analyzing, for instance, the relationship between exposure to a particular chemical substance and a health problem, or between consumption of a food and defined positive health effects (Luján and Todt 2021).

As a general rule, health claims fall into the category of benefit assessment. But we have to distinguish different types of health claims. We can distinguish between health claims directly aimed at exposing health benefits and health claims aimed at showing that a food (or ingredient) serves to reduce the risk of a disease. This second case is the one that Jukola (2019) focuses on. She examines the assessment of evidence in the case of nutrition guidelines at the population level, where the main issue is to understand how a certain

type of diet can prevent chronic diseases. To the contrary, we emphasize direct benefits that concern those assessments of evidence focused on enhanced function claims on foods or their ingredients.

In 1995, the European Union (EU) and the European ILSI (International Life Sciences Institute) started the FUFOSE project (Functional Food Science in Europe) in order to create a new approach to evaluating scientific evidence necessary for the development of functional foods (Aggett et al. 2010).[8] Its primary aim was to adequately characterize the notion of health claims, whether conceived as enhanced function claims or as reduction of disease risk claims (Agarwal et al. 2006). In order to implement the conclusions and principles of FUFOSE, in 2005 the PASSCLAIM Project (Process for the Assessment of Scientific Support for Claims on Foods) was started, whose final version defined several criteria to substantiate a health claim. The EU's Health Claim Regulation (European Parliament and Council 2006) has since sought to make it easier for the consumer to choose scientifically substantiated claims and to communicate benefits from food, provided that the claim is actually substantiated scientifically (Art. 6.1) (Asp and Bryngelsson 2008).

The use of scientific evidence in regulation follows criteria defined by EFSA's relevant scientific committee, the NDA Panel (NDA Panel EFSA, 2011).[9] Taking these criteria

[8] On controversies with respect to ILSI, see https://corporateeurope.org/sites/default/files/ilsi-article-final.pdf.

[9] The scientific data have been structured according to their relevance in this order: (1) human intervention studies, (2) human observational studies, (3) other human studies, and (4) non-human

into account, in the following we review a number of issues that question the idea of being able to generate data on causal relationships between intake and outcome without recurring to expert judgment.

[Food matrix] In order to verify the effects of a nutrient or food component, it is necessary to characterize certain *background conditions* that reflect the food matrix (background diet), as well as the dietary context of the intake of the nutrient under study. The food matrix can decisively influence the release and activity of the ingredient in question (PASSCLAIM 2005). It is therefore important to calibrate this influence. For example, using water-free *in vitro* food systems allows to compare the release of the components of the different dietary matrices (Rietjens and Alink 2006). This shows that in nutrition science any RCT data used to ascertain effects from foods are typically more complex than those used in pharmacology, because certain features of nutrition set it apart from pharmacology—low effective doses, interaction among all of the ingredients involved, interplay between the ingredients under study and the food matrix, as well as relatively long latency periods (Hendrickx 2013).

[Extrapolation] In nutrition science, as a DOSS, few results are extensible to all foods and food components. A health claim obtained with a particular diet or food matrix cannot always be *extrapolated* to a second product with the same component in a different food matrix (the nutrient and non-nutrient components of foods). Applying the use of an existing health claim to a product with another composition requires evidence of the invariance of functional efficacy,

---

data as evidence for supporting claims (EC, 2008; see also EFSA (NDA Panel, Regulation (EC) No 1924/2006, 2011, p. 2021).

which in turn may require a separate substantiation of the reported effect for each individual product.

[Consistency] Along with the characterization of the background diet and other relevant aspects of the study groups' lifestyle (PASSCLAIM 2005, p. 17)—where the use of independent markers of intake and exposure helps increase the quality of data on dietary intake—,[10] another important point is the consistency between the amount of food and the intended pattern of consumption: the amount should approximate or coincide with its intended use, the form and the frequency of the intake. Where dose-response studies are performed (concentration-effect), the dose range usually includes the amount of food expected to be consumed (Neale and Tapsell 2019, p. 3; Jacobs and Tapsell 2013). In the food supplied to the control group of a trial the ingredients under study will be either absent or present in a known and significantly different—usually much lower—concentration with respect to the amount present in the tested food. The analytic procedure under certain circumstances may also have subjective effects that might not actually be related to specific effects of the tested substance itself. These effects can be placebo or nocebo phenomena, and may be considered in the study design.

[Accuracy] Methodological accuracy requires that the exposure and monitoring of the food are adequate. This involves a sufficiently large observation period (period of intake) in order to show that the expected effect exists and,

[10] Whenever experimenters do not have valid markers of exposure at hand, intakes can be calculated according to the amount and composition of the food consumed, which requires not only that subjects reliably report on their food intake, but also that reliable information about the composition of food for that population is available from a database.

if necessary, that it can be maintained over time. Since the period of observation may range from a few seconds to several weeks (sterol-induced changes in cholesterol metabolism) or years (changes in bone density due to calcium), meaning that studies might in practice be extremely challenging or even impossible, it can be necessary to recur to alternative approaches to assess reported benefits, such as the use of biomarkers.

In view of the above, and given that some of the basic characteristics of RCTs in pharmacology (blinding of participants and investigators, use of appropriate controls, choosing control groups, time until a relevant outcome is produced, etc.) turn out to be highly problematic in the field of nutrition, we can conclude that many RCTs— particularly long-term ones—are difficult or even impractical in nutrition research (Neale and Tapsell 2019, p. 6). This implies two conclusions: 1) it may be *necessary* to recur to methodologies other than RCTs to generate data, like cohort studies—research used to investigate the causes of disease, and to establish links between risk factors and health outcomes—and 2) expert knowledge becomes crucially important. In other words, the methodological limitations of RCTs and the need for data from other, non-RCT sources in nutrition research mean that experts are even more important in nutrition (and related regulatory decisions) than in the field of pharmacology. We agree with Schwingshackl et al. (2016) and Neale and Tapsell (2019, p. 6) that one partial solution to this difficulty may be scoring systems that allow to evaluate the quality of the evidence generated by RCTs (like, for instance, NutriGrade). Good *systematic reviews* could take advantage of a framework like this, although institutions and companies should finance this type of evidence production control, at least if we do not want misinformation to

constantly sneak into control programs quality (See Schwingshackl et al. 2016).

5. Conclusions

Randomization requires the concurrence of expert judgment because, as we claim, the "automatic objectivity" ascribed to RCTs is insufficient for taking valid regulatory decisions. Assessments, in order to be useful, have to rely on the assessment of data from diverse sources of evidence. The latter implies the concurrence of a variety of scientific methodologies, which in practice makes expert judgment indispensable. Expert judgment, in other words, is a necessary complement to "objective" RCT data. Even more, expert intervention may also be required when designing and executing RCTs. For instance, it is unlikely that RCT evidence can be adequately generated with respect to the role a nutrient plays in endpoints of non-indexed diseases without taking account of expert judgment. Much of the evidence regarding nutrients and non-indexed diseases will in the end have to recur to a wider variety of methods, among them mechanistic and observational studies.

In this paper we have analyzed whether obtainment of scientific evidence by way of RCTs is the most adequate (or even the only valid) procedure for regulatory decision making in the field of nutrition. It is clear that minimization of bias is a key characteristic of RCTs. But this is not sufficient for excluding other evidence from decision making. It is necessary to take into account other strategies for modeling evidence (Zeiss and van Egmond 2014), since researchers in most cases are interested not only in predicting the effects of a nutrient, but also in *explaining* it, or *applying* it under conditions of high uncertainty. Therefore,

pharmacology-style RCTs, given their methodological limitations, are not adequate as *exclusive* tools for generating scientific evidence in nutrition, since the generation of regulation-relevant data in this field tends to be more complex than in pharmacology. The RCT methodology needs to be complemented and 'interpreted' (*shaped*) with empirical nutritional information (food details, data on the potential consumer, etc.), just as in physical-mathematical models the initial conditions, the boundary conditions and other elements are taken into account in order to obtain unbiased and objective outcomes (Garza et al. 2019).

In short, we have examined the nature and quality of evidence generated by RCTs to enable scientists and regulators to judge and manage decisions about the acceptability and truthfulness of health claims. Our analysis shows that both RCTs and expert knowledge are *jointly* necessary in data generation for regulation and decision making. As we have argued, even in pharmacology RCTs depend on expert input. In section 4 we have seen that as far as data generation is concerned, nutrition is more complex than pharmacology, implying that RCTs are more difficult to effectively design and execute. What the latter means is that in nutrition and health claim regulation, expert knowledge is even more important than in pharmacology.

References

ACHINSTEIN, P. *The Book of Evidence*. Oxford: Oxford University Press, 2001.

AGARWAL, S., HORDVIK, S., MORAR, S. "Nutritional claims for functional foods and supplements". *Toxicology*, 221, pp. 44-49, 2006.

AGGETT, P., ANTOINE, J.-M., DE VRIES, J., GALLAGHER, A., HENDRIKS, H., KOZIANOWSKI, G., MEIJER, G., RICHARDSON, D., RONDEAU, V., TWEEDIE, G., WELCH, R. (ILSI), "Beyond PASSCLAIM – Guidance to substantiate health claims on foods". *ILSI Europe Report Series*, pp. 1-24, 2010.

ASP, N.-G., BRYNGELSSON, S. "Health claims in Europe: new legislation and PASSCLAIM for substantiation". *Journal of Nutrition*, 138, pp. 1210S–1215S, 2008.

BLUMBERG, J., HEANEY, R.P., HUNCHAREK, M., SCHOLL, T., STAMPFER, M., VIETH, R., WEAVER, C.M., ZEISEL, S.H. "Evidence-based criteria in the nutritional context". *Nutrition Reviews*, 68, pp. 478-484, 2010.

BOWEN, S., ZWI, A.B. "Pathways to 'Evidence-Informed' Policy and Practice: A Framework for Action". *PloS Medicine*, 2, pp. 0600-0605, 2005.

CARTWRIGHT, N. "What are randomised controlled trials good for?". *Philosophical Studies*, 147, pp. 59-70, 2010.

_____. *Evidence: For Policy, and Wheresoever Rigor is a Must*. London: LSE, 2015.

_____., HARDIE, J. *Evidence-Based Policy: A Practical Guide to Doing It Better*. Oxford: Oxford University Press, 2012.

DE BOER, A., VOS, E., BAST, A. "Implementation of the nutrition and health claim regulation – The case of antioxidants". *Regulatory Toxicology and Pharmacology*, 68, pp. 475-487, 2014.

DOUGLAS, H. "The Irreducible Complexity of Objectivity". *Synthese* 138, pp. 453-473, 2004.

EFSA. Scientific and technical guidance for the preparation and presentation of an application for authorisation of a health claim (Revision 1). *EFSA Journal*, 9, 2170, pp.1–36, 2011.

EUROPEAN PARLIAMENT AND COUNCIL. 2006. Regulation (EC) 1924/2006 of the European Parliament and of the Council of 20 December 2006 on nutrition and health claims made on foods. *Official Journal of the European Union* L.404: 9-25.

EU, 2002. "Regulation (EC) No 178/2002 of the European Parliament and of the Council of 28 January 2002 laying down the general principles and requirements of food law, establishing the European Food Safety Authority and laying down procedures in matters of food safety." *Chapter 11 'General Food Law', Article 6 'Risk analysis' and article 14 'Food safety requirements'. Official Journal of the European Communities* (1.2.2002), L31, 34 pp. (Available at: http://eur-lex.europa.eu/legal-content/EN/TXT/PDF/?uri=CELEX:32002R0178&from=en. Accessed on May 9th, 2020).

GARZA, C., STOVER, P.J., OHLHORST, S.D., FIELD, M.S., STEINBROOK, R., ROWE, S., WOTEKI, C., CAMPBELL, E. "Best practices in nutrition science to earn and keep the public's trust". *The American Journal of Clinical Nutrition*, 109, pp. 225-243, 2019.

GOODMAN, K. *Ethics and Evidence-Based Medicine: Fallibility and Responsibility in Clinical Science*. Cambridge: Cambridge University Press, 2003.

HACKING, I. *Representing and Intervening*. Cambridge: Cambridge University Press, 1983.

HAYWARD, R.A., KRUMHOLZ, H.M. "Three reasons to abandon low-density lipoprotein targets". *Circulation*, 5, pp. 2-5, 2012.

HEANEY, R.P. "Nutrients, Endpoints, and the Problem of Proof". *The Journal of Nutrition*, 138, pp. 1591-1595, 2008.

HENDRICKX, K. "Rivaling evidence-bases and politics in regulatory science". *Food, Science & Law*, 4, 2013 [http://hdl.handle.net/2268/162196] (Access: May 29, 2021).

HSIANG-K., C., SZU-T., C., MILLSTEIN, R.L. *Mechanisms and Causality in Biology and Economics*. Dordrecht, Springer, 2013.

ILLARI, P., RUSSO, F. *Causality: Philosophical Theory Meets Scientific Practice*. Oxford: Oxford University Press, 2014.

JACOBS, D.R., TAPSELL, L.C. "Food synergy: the key to a healthy diet". *Proceeding of the Nutrition Society*, 72, pp. 200-206, 2013.

JIMÉNEZ-BUEDO, M., MILLER, L.M. "Why a Trade-Off? The Relationship between the External and Internal Validity of Experiments". *Theoria* 69, pp. 301-321, 2010.

JUKOLA, S. "On the evidentiary standards for nutrition advice". *Studies in History and Philosophy of Biological and Biomedical Sciences* 73, pp. 1-9, 2019.

KOSKINEN, I. "Defending a risk account of scientific objectivity". *The British Journal for the Philosophy of Science*, 71, pp. 1187-1207, 2020.

LÄHTEENMÄKI, L. "Claiming health in food products". *Food Quality and Preference*, 27, pp. 196-201, 2013.

LAWRENCE, G.D. "Dietary fats and health: dietary recommendations in the context of scientific evidence". *Advances in Nutrition*, 4, pp. 294-302, 2013.

LUJÁN, J.L., TODT, O. "Evidence based methodology: a naturalistic analysis of epistemic policies in regulatory science". *European Journal for Philosophy of Science* 11**,** pp. 26-44, 2021.

NEALE, E.P., TAPSELL, L.C. "Perspective: The Evidence-Based Framework in Nutrition and Dietetics: Implementation, Challenges, and Future Directions". *Advances in Nutrition*, 10, pp. 1-8, 2019.

NHCR-Regulation (EC). 2006. No 1924/2006 of the European Parliament and of the Council of 20 December 2006 on nutrition and health claims made on foods (17 pp.) (Available at: http://eur-lex.europa.eu/legal-content/en/ALL/?uri=CELEX%3A32006R1924. Accessed on December 2nd, 2020).

PAPINEAU, D. "The Virtues of Randomization". *The British Journal for the Philosophy of Science*, 45, pp. 437-450, 1994.

PASSCLAIM, "Process for the Assessment of Scientific Support for Claims on Foods: Consensus on Criteria". *European Journal of Nutrition*, 44, §§ I/1-I/2, 2005.

REISS, J. "A pragmatist theory of evidence". *Philosophy of Science*, 82, pp. 341–362, 2015.

RIETJENS, I., ALINK, G. "Future of toxicology–low-dose toxicology and risk– benefit analysis". *Chemical Research in Toxicology*, 19, pp. 977–981, 2006.

RUBIN, D.B. "Estimating Causal Effects of Treatments in Randomized and Nonrandomized Studies". *Journal of Educational Psychology*, 66, pp. 688-701, 1974.

SACKETT, D.L., ROSENBERG, W.M.C., GRAY, J.A.M., HAYNES, R.B., RICHARDSON, W.S. "Evidence based medicine: what it is and what it isn't". *British Medical Journal*, 312, pp. 71-72, 1996.

SCHWINGSHACKL, L., KNÜPPEL, S., SCHWEDHELM, C., HOFFMAN, G., MISSBACH, B., STELMACH-MARDAS, M., DIETRICH, S., EICHELMANN, F., KONTOPANTEILS, E., IQBAL, K. "Perspective: NutriGrade: a scoring system to assess and judge the meta-evidence of randomized controlled trials and cohort studies in nutrition research". *Advances in Nutrition*, 7, pp. 994-1004, 2016.

STEGENGA, J. "Measuring effectiveness". *Studies in History and Philosophy of Biological and Biomedical Sciences*, 54, pp. 62-71, 2015.

STEEL, D. *Philosophy and the precautionary principle: Science, evidence, and environmental policy*. Cambridge: Cambridge University Press, 2015.

TAPSELL, L.C. "Evidence for health claims: a perspective from the Australia–New Zealand region," *The Journal of Nutrition*, 138, pp. S1206-S1209, 2008.

TEIRA, D. "On the impartiality of British trials". *Studies in History and Philosophy of Biological and Biomedical Sciences*, 44, pp. 412-418, 2013.

TEIRA, D., REISS, J. "Causality, Impartiality and Evidence-Based Policy". In C. Hsiang-K., C. Szu-T. and R. L. Millstein (eds.) (2013), p. 207-224.

TRUSSWELL, S. "Levels and kinds of evidence for public-health nutrition". *The Lancet*, 357, pp. 1061-1062, 2001.

URBACH, P. "Randomization and the design of experiments". *Philosophy of Science*, 52, pp. 256-273, 1985.

VAN BAALEN, S., BOON, M. "An epistemological shift: from evidence-based medicine to epistemological responsibility". *Journal of Evaluation in Clinical Practice*, 21, pp. 433-439, 2015.

WORRALL, J. "Evidence: philosophy of science meets medicine". *Journal of Evaluation in Clinical Practice*, 16, pp. 356-362, 2010.

ZEISS, R., VAN EGMOND, S. "Dissolving Decision Making? Models and Their Roles in Decision-Making Processes and Policy at Large". *Science in Context*, 27, pp. 631–657, 2014.