

# Hill's "Biological Plausibility" Criterion: Integration of Data from Various Disciplines for Epidemiology and Radiation Epidemiology<sup>1</sup>

A. N. Koterov<sup>a, \*</sup>, L. N. Ushenkova<sup>a</sup>, and A. P. Biryukov<sup>a</sup>

<sup>a</sup> *Burnasyan Federal Medical Biophysical Center of the Federal Medical Biological Agency, Moscow, Russia*

\**e-mail: govorilga@inbox.ru*

Received September 16, 2019; revised September 29, 2019; accepted November 20, 2019

**Abstract**—In this review, various aspects of the causality criterion "Biological Plausibility," which is sometimes replaced by the criterion "Coherence" (consistency with well-known medical and biological knowledge), are considered. The importance of the criterion for epidemiological evidence of causation, especially for disciplines such as ecology, toxicology, and carcinogenesis, in which there are difficulties not only to perform experiments, but even to observe the effect, is noted. Only statistical approaches in epidemiology are incapable of proving the true causality for association (possibly the effect of chances, confounders, biases, and reverse causation). Without knowledge of the biological mechanism and a plausible model, such a relationship (especially for weak associations) cannot be regarded as confirmation of the true causation of the effect of the factor. The essence of the criterion is the integration of data from various biomedical disciplines, including molecular and experiments on animals and in vitro. There are three (D.L. Weed and S.D. Hursting, from 1998) and four (M. Susser from 1977 and 1986) levels of attaining biological plausibility and coherence with the existing knowledge. The methodologies for integrating data from various disciplines through Bayesian meta-analysis, based on Weight of Evidence (WoE) and teleoanalysis, are considered. The latter is a combination of data from different types of studies to quantify the causal relationship between two such associations, each of which can be proved. However, it is difficult for several reasons, including an ethical plan, to determine the relationship between the causality of the first and the final effect of the second. The approach by teleoanalysis seems doubtful. Despite the need for the "Biological Plausibility" criterion, it, similarly to almost all of Hill's criteria (except for "Temporality"), is neither necessary nor sufficient for evidence. Examples are presented (including the effects of radiation) that show that, first, "Biological Plausibility" depends upon the biological knowledge of the day and, second, there are real though seemingly implausible associations, and vice versa. This is basis for criticism by some authors (A.R. Feinstein; K.J. Rothman and S. Greenland; B.G. Charlton; K. Goodman and C.V. Phillips) both specifically for the criterion "Biological Plausibility" and for the entire inductive approach based on causal criteria. However, the "Biological Plausibility" criterion remains important for proving causality in epidemiological studies, especially for those disciplines that public health relies on in making preventive decisions and developing safety standards.

**Keywords:** causality criterion "Biological Plausibility," weight of evidence, integration of data from biomedical disciplines, Bayesian meta-analysis, teleoanalysis

**DOI:** 10.1134/S1062359021110054

To date, our cycle of publications in the journal "Radiation Biology: Radioecology," devoted to causality issues in observational disciplines, including the radiation profile, has expanded from the originally planned three reviews to a minimum of four. Two reviews have already been published [1, 2], and two are planned. They are devoted to causation models and description of the historical origins, essence, limitations, breadth of application, and radiation aspect of the guidelines/criteria by which causality is estab-

lished [1, 2], starting from Henle–Koch's postulates for infectious diseases (J. Henle and R. Koch, 1877–1893, Germany) with their subsequent improvements and ending with Hill's criteria (A.B. Hill, 1965, Great Britain; M. Susser, 1973, United States), Evans (A.S. Evans, 1976, United States), and other authors who are less well-known today [1, 2] (and subsequent communications).

The material published on the topic since the early 1950s (almost entirely from the United States [2]), of both epidemiological and scientific–philosophical nature, including articles and voluminous Western

<sup>1</sup> Published in the author's version.

textbooks on epidemiology and carcinogenesis (for the hundreds of sources and more than 40 textbooks used by us, see [3]), is so ample that, in addition to the unifying Reports nos. 1–4 per se, there was a need for separate detailed publications on the most important criteria. Two reviews on the “Strength of Association” [4, 5] and one review on “Temporality” with description of evidence of reverse causality [3] have already been published. This work is devoted to another criterion, “Biological Plausibility,” because the corresponding material in terms of its volume, importance, and relevance, including for experimental radiobiology, is difficult to reflect only in a separate fragment of the article.

Thus, reviews [3–5] and the review presented below are some detailed preambles to Reports 3 and 4<sup>2</sup>.

The importance of the problems under consideration lies in the fact that, in all descriptive disciplines, which includes not only natural sciences but also retrospective (history), socioeconomic, psychological, etc., disciplines (see [3, 4]), evidence can be based on identifying statistically significant associations between cause and effect, between exposure and effect, between the characteristics of a group and its subsequent behavioral features, etc. In experimental disciplines, where it is possible to determine the conditions of the experiment, obtaining such evidence is quite clear (an approach is called experimental when at least one out of many variable factors can be controlled [8]). *Experimental* identification of a statistically significant association or correlation is the *final* stage of evidence [9] (which also applies to randomized controlled trials (RCT) in medicine [8, 10]). However, for descriptive/observational disciplines, association does not mean causality, whatever the statistical significance of the correlation is [1–7, 9–12].

The identified association in observational studies in the absence of control over the variant(s) can be explained by the following unaccounted factors [13] (and other sources; see [1–5]):

- chance;
- interfering factors (“third factor,” “confounder”);
- systematic errors (“bias”);

<sup>2</sup> It would be useful to present the material on the problem of causality and methodologies for its establishment (criteria, etc.) in the form of a monograph, similarly to the first edition of this kind by M. Susser in 1973 [6] and, for example, the collection of 2015 (497 pp.) on causality in ecoepidemiology edited by the leading authors from the US Environmental Protection Agency (USEPA) S.B. Norton, S.M. Cormier, and G.W. Suter II [7]. A monograph on this topic in Russian may be relevant due to weak corresponding concepts in Russian natural science, medical, and epidemiological areas, as well as in everyday scientific and everyday life in Russia (see [1–5]). However, there is no incentive to create such a monograph.

- reverse causation.

All these factors with numerous examples were considered by us earlier [1–5] (except for the biases, a list of which can be found in many textbooks, and more than three dozen of which are given in the Oxford Epidemiological Dictionary [14]).

In this regard, for epidemiology, the establishment of a statistically significant association between two phenomena, in contrast to experimental disciplines, is only the very first, *initial* stage of proving causality [9] (and other works, see [3, 4]).

To confirm the causality of association, following the Henle–Koch postulates of the 19th century for infectious diseases, in the 1950s–1970s in epidemiology a number of “precautions,” “viewpoints,” “guidelines,” “judgments,” “criteria,” “postulates,” etc., were developed (these are essentially synonyms; see sources in [1–3]) to assess the causality of chronic, noninfectious pathologies. The most well-known are Hill’s nine causality criteria, eight of which this authoritative English statistician in the field of medicine in the past only collected together, having taken them from other authors [11]. Nevertheless, now the criteria of causality in epidemiology [1–5, 14] and even in ecoepidemiology [7, 15, 16] are almost always called Hill’s criteria or Hill’s guidelines.

Hill himself emphasized that, with the exception of “Temporality” (the exposure should precede the effect) [3], these are not strict rules-criteria but some principles, guidelines for assessing the degree of probability that the association is causal (“viewpoints”) [11].

According to the analysis of authoritative authors (Weed and Gorelic [17], etc.), as well as the data of our study of publications for 2013–2019 (we plan to consider these materials in more detail in Reports 3 and 4), some criteria seem to be the most important and are used more often.<sup>3</sup> Therefore, following, as already mentioned, the above-mentioned preamble reviews [3–5], devoted to two main criteria, in this work we consider the third of them, “Biological Plausibility.”

#### HISTORY OF THE “BIOLOGICAL PLAUSIBILITY” CRITERION: OVERLAPPING MEANINGS WITH “COHERENCE”

A suitable point, which was later called a “postulate,” first appeared in the work by E.L. Wynder in 1956 [18], containing the rules for establishing the

<sup>3</sup> The currently active D.L. Weed (United States) for decades has paid great attention to the theory and practice of causation (many publications and synthetic studies), as a consulting expert on problems at the intersection of biomedical disciplines, law, commerce, and public policy (see also [2]). In our works on causality [1–4], Weed’s publications are cited probably more often than the works of other authors.

causality of cancer: the agent must be shown to be carcinogenic in some animal species (not essential).<sup>4</sup> But in 1961 Wynder and Day [19] removed this point from the list when the postulates were extended to all chronic pathologies (i.e., the author seemed to change his mind).

However, A.M. Lilienfeld in 1957 [20] developed this point. The term "Biological Plausibility [of a causal hypothesis]" per se was introduced. The point was ensured by experiments on animals and had gradations of magnitude. Then, A.M. Lilienfeld in 1959 [21] wrote about the biological plausibility of an association that depends on our basic knowledge of the biology of these specific pathologies.<sup>5</sup> In 1960, P.E. Sartwell [22], among his five "points," after the strength of the association, its reproducibility, the correlation itself (for some reason, in third place), and the temporal dependence (for some reason, in fourth place), called the last one "The Biological Reasonableness [of the association]." Pondering the logical order of the point was, probably, not the purpose of the short paper by Sartwell in 1960 [22]; only their meaning was important.

In the Report of the Chief Physician of the United States on the consequences of smoking, published in 1964 [23], this point is absent; it was replaced by "Coherence," agreement with the known facts of the natural history and biology of the disease.<sup>6</sup> It can be seen that this point almost literally coincides with the "Biological Plausibility" from Lilienfeld in 1959 [21].

In 1965, Hill [11] supplemented "Biological Plausibility" (probably from the works by Lilienfeld from 1957 and 1959 [20, 21]) with the point "Coherence," probably from the Report of 1964 [23] (there are no references in [11] to anyone [2]). However, Hill's first criterion received a slightly different meaning, namely,

<sup>4</sup> "The agent should be shown to be carcinogenic to some animal species (not obligatory)" [18].

<sup>5</sup> "... biological plausibility of the association, which is dependent upon our general knowledge of the biology of these specific diseases" [21]. The conceptual combination "biologically implausible" took place much earlier, for example, in the decision of the British public health authority in 1854, according to which epidemiological evidence of cholera infection through London water (researcher J. Snow) was not supported by significant laboratory evidence (cited from [22]).

<sup>6</sup> The 1964 report of the US Surgeon General (Chief Physician [24]) on the consequences of smoking [23] of wide international significance [23] is the first key milestone in the formalization of methods for proving the causality of chronic pathologies, including cancer. It is regarded as the official date of the final establishment of the association between smoking and lung cancer (see details in [2]).

the actual plausibility of dependence in a broad biological sense (not only in relation to a specific pathology).

M. Susser, at least from 1977 [25], developed the "Coherence" criterion, which essentially included "Biological Plausibility".<sup>7</sup> In the widely cited textbook "Pharmacoepidemiology" (edited in 2000), "Coherence" and "Biological Plausibility" also completely coincide [27].

Thus, the "Coherence" criterion has supplanted "Biological Plausibility" in a number of sources. It should also be noted that the meaning of "Biological Plausibility" sometimes coincides with what is now understood under the criterion "Experiment" [16, 27].

In Hill's criteria modified for ecoepidemiology, "Biological Plausibility" and "Coherence" are replaced by "Biological Concordance" [16].

According to the aforementioned study by Weed and Gorelic [17], among the 14 reviews on the topic, the "Biological Plausibility" criterion was mentioned in eight sources (fourth place among the nine criteria). In our analysis of publications for 2013–2019 (35 papers in which Hill's criteria were used as a methodology), this statement in terms of its use was in third place (in first place, "Strength of Association"; in the second place, with the same results, were "Temporality," "Consistency," and "Biological Gradient," i.e., the "dose–effect" relationship).

#### ESSENCE OF THE CRITERION: INTEGRATION OF DATA FROM VARIOUS DISCIPLINES

The topic of introducing data from biology, medicine ("biomedical knowledge" [28]), toxicology, pharmacology, ecology, and other disciplines [2, 11, 15, 16, 23–59] into epidemiological evidence, as well as the reverse approach—verification biological facts by epidemiological patterns (this also takes place) [52]—is very broad and, indeed, deserves a separate review, which was carried out before us, though long ago by Weed and Hursting in 1998 [29].<sup>8</sup>

"Epidemiology, molecular pathology (including chemistry, biochemistry, molecular biology, molecular virology, molecular genetics, epigenetics, genomics, proteomics, and other molecular approaches), and in vitro animal and cell experimentation should be

<sup>7</sup> In the aforementioned monograph by Susser from 1973 [6], which is unavailable, the point "biological theory and experiment," according to [26], was also considered and called "the final criterion for causation."

<sup>8</sup> All relevant references used in this review are provided.

considered an important integration of evidence into determination of carcinogenic effects for humans” (Cancer Institute, United States, 2012) [53].<sup>9</sup> From the first paragraph, we can add here medicine, toxicology, pharmacology, and ecology.

This list shows that data from almost every area of biomedical and molecular–cellular disciplines can make an important contribution to the search for evidence for epidemiological research of any practical importance. This, clearly, gives practical significance to the research of any seeming fundamental and theoretical nature.

The essence of the criterion is generally clear; it is verbosely, but vaguely, explained in numerous works (“No clear rule-of-thumb exists for this criterion” [51]). This is probably why the available material sometimes resembles a “stream of consciousness” [29, 54, 60, 61].<sup>10</sup> Such material is difficult to systematize and categorize according to meanings and levels of significance.

*Compliance with general scientific knowledge.* The latter include scientific facts and laws that relate to the alleged causation (general scientific) [55]. The criterion refers to the *scientific* plausibility of the association [43].

*Support by in vitro and animal laboratory experiments* [27, 62, 63]. A question arises as to what extent the models obtained for cells and animals are applicable to humans? If the effect is observed in rats or monkeys, then how is it known that it will be observed in humans? Conversely, just because the effect is not found in animals, does it follow that it cannot occur in humans [62]? (The deplorable example of thalidomide taken by pregnant women is given below.)

“Despite these uncertainties, experience has shown that animal evidence can matter and, where it exists, should be taken into account” (2019 handbook) [62].<sup>11</sup> For example, teratogens in humans were also found to be teratogens in at least one animal species and exhibit some toxicity in the majority of the species tested. However, as a rule, the extrapolation of effects between species requires knowledge of their physiolo-

gy [15] (the issues of the applicability of the laboratory approach, within the framework of the “Experiment” criterion, are planned to be considered in our further works).

*The presence of biological and social models and mechanisms that explain the relationship* [13, 34, 38, 62, 64, 65]. They are considered at different levels of biological and social organization, from molecular–cellular to population (for example, at the level of behavioral reactions contributing to the carcinogenicity of the factor (“social”)) [29]. “A causal relationship is considered biologically plausible when evidence of a probable causal mechanism is found in the scientific literature” [66] (based on [66], the idea arises that it is not necessary to carry out experiments for confirmation, because synthetic studies are sufficient).<sup>12</sup> T. Hartung in 2007 called it “mechanistic validation” [56].

Biological plausibility reflects the consistency/inconsistency of the theory explaining how or why exposure causes a disease, with other known mechanisms of causation of that disease [31]. For example, was it shown that the agent or metabolite can reach the target organ at all? [67]. “Scientific or pathophysiological theory” [27]. “Biological common sense” [15]. In medicine, it is customary to begin the study of a condition, for example, a rare congenital disease, with observations and epidemiological data, and end it with the study of the mechanisms of its inheritance and the essence of metabolic “break-down” [24].

International and national agencies have compiled lists of agents and effects that are considered carcinogenic to humans. The criteria are, in addition to the actual epidemiological studies, data from animal experiments and ideas about the mechanisms. However, data on the mechanism alone can rarely be considered adequate for establishing the carcinogenicity of an agent in relation to humans, despite all the advances in understanding the molecular basis of this process [47] and all the “precautionary principles” [1]. According to [68], to make a conclusion about causality, in addition to the existence of a plausible mechanism, epidemiological data on the frequency of the effect are, nevertheless, required. Likewise, review [54] indicates that mutual support of mechanisms and dependencies is required to establish causal relationships. The same principle is followed by the International Agency for Research on Cancer (IARC) for carcinogenic factors: the assessment of causality depends on the presence of a plausible mechanism and proba-

<sup>9</sup> “Epidemiology, molecular pathology (including chemistry, biochemistry, molecular biology, molecular virology, molecular genetics, epigenetics, genomics, proteomics, and other molecular-based approaches), and animal and tissue culture experiments should all be seen as important integrating evidence in the determination of human carcinogenicity” (National Cancer Institute (United States), workshop 2003) [53].

<sup>10</sup> Particularly impressive is the review by Ward [61], where 19 pages of the PDF text (preceding the list of references) have the headings “Abstract,” “Introduction,” “Analysis,” and “Conclusion.”

<sup>11</sup> “Despite these uncertainties, experience has shown that animal evidence can be relevant, and where it exists it should be taken into account” [62].

<sup>12</sup> “The causal relationship was considered biologically plausible when evidence for a probable cause–effect mechanism was found in the scientific literature” [66].

bilistic data, i.e., on the increased incidence of cancer in the population or on the relative risk [54, 57].<sup>13</sup>

*Three levels of achievement of biological plausibility according to D.L. Weed* (cited in the thematic review by Weed and Hursting [29] (see also [51]):

(1) When a reasonable mechanism for association can be assumed, but there is no biological evidence (the beginning was made, according to [29], in [11] and repeated by subsequent authors, for example, [13]).

(2) When facts from the field of molecular biology and molecular epidemiology can be added to the mechanism (we add the following: including the data on "surrogate EndPoints," i.e., biomarkers [7]). In [29], the IARC monograph from 1990 is given as an illustration; however, there are similar constructions in later documents of this organization. Particular attention is given to molecular and cellular indicators: adducts and DNA lesions, their repair, proliferation disorders, level of cell death, changes in intercellular connections, etc. (IARC-2006) [57].

(3) If there is sufficient evidence of how the causative factor affects the known mechanism of pathology [69, 70]. This is the most rigorous of the three approaches to biological plausibility [29]. Weed [39] indicates that, although it is difficult to propose a rule of thumb for determining the biological plausibility, if the factor of interest is the key factor in a biological mechanism or in its pathways, then it is more likely that the dependence will be causal. Mechanisms or pathways are demonstrated using laboratory experiments and molecular epidemiology [39].

For these three levels, Weed and Hursting [29] developed scholastic reasoning, not devoid of theoretical meaning but difficult to translate into practice. They point out that it is unclear how much evidence and what evidence will turn the "suggested" (Hill and other authors) [11, 13] or "hypothesized" (Susser [71]) mechanism into a "coherent" mechanism (Report of the Chief Physician of the United States on the consequences of smoking from 1964) [23], i.e., the one that not only "makes sense" [72] but is also defined by our detailed understanding of each step in the chain of events [23].<sup>14</sup> The question also arises as to what is

required to assert that we "know" the mechanism [69, 70]? The fact is that an explanation (i.e., a mechanism) is understood in a broad sense, because it is difficult to define what a "good explanation" is and how to take into account the "success" of mechanisms [54].

After all, it is all a matter of judgment whether there is enough evidence to rule out alternative explanations ("Weight of Evidence," WoE [16, 34, 38, 40, 54, 56, 67]) [30]. Earlier, the Committee on Diet and Health of the National Research Council's Commission on Life Sciences for a complex of six of Hill's criteria (strength of association, dose dependence, temporal dependence, consistency of association, specificity, and biological plausibility) gave equal weight to all criteria except plausibility, which was downgraded as this provision depends on subjective interpretation [73], which later was considered "very subjective" [74].

*Four levels of acquisition of biological plausibility according to M. Susser.* This author, as mentioned above, had been developing the concept of the "Coherence" criterion, including "Biological Plausibility," since 1977 [25] (if not earlier; see note 6), and in 1986 [36, 37] he outlined his concepts in detail. The criterion reflects the pre-existing theory and knowledge (the "Coherence" of other authors [11, 23]) and is interpreted broadly, including the following elements of coherence: (1) with theoretical plausibility (also named in [54]), (2) with facts, (3) with biological knowledge (i.e., "Biological Plausibility"), and (4) with statistical patterns including the dose-effect relationship (we plan to consider the set of criteria proposed by Susser in more detail in Report 4).

#### THE IMPORTANCE OF INTEGRATION OF BIOLOGICAL AND OTHER DISCIPLINES WITH EPIDEMIOLOGY (RADIATION EPIDEMIOLOGY): QUOTES

It is appropriate to provide notes of other authors (mainly from significant sources) reflecting this provision. They can be useful for experimenters for substantiation (for this purpose, the originals are also presented in the section "Notes").

— Advances in the biological sciences and their integration into health care through molecular epidemiology make one causal criterion, "Biological Plausibility" (sometimes called "Biological Coherence"), an increasingly important consideration in causal inference (1998) [29].<sup>15</sup>

<sup>13</sup>The provision from [54, 57, 68], according to which, for the conclusion about causality, in addition to the very fact of association and time dependence, only "Biological Plausibility" in the form of a mechanism shown experimentally is sufficient and can be used with the necessary conjuncture (significant sources). The latter, however, can be done only with caution, since, as will be shown below, there are enough exceptions and the very concept of "mechanism" is limited by the framework of the achieved level of knowledge.

<sup>14</sup>Article [72] deals with the issue of which epidemiological data will be published by editors. It is unlikely that there will be information about an effect that is both weak and biologically implausible. However, a report of a small effect that has biological sense or a large effect that does not have it may be published.

<sup>15</sup>"Advances in the biological sciences and their integration with public health science in molecular epidemiology make one causal criterion, biological plausibility (sometimes called biological coherence), an increasingly important consideration in causal inference" [29].

— Biological plausibility represents fundamental concepts of data integration—the criterion suggests that epidemiology and biology must interact (2015) [34].<sup>16</sup>

— Correlations and data must have biological and epidemiological sense (1983) [75].<sup>17</sup> Causality must have biological meaning (2016) [33].<sup>18</sup>

— Biological plausibility is a criterion where biology and epidemiology merge (2016) [46].<sup>19</sup>

— Demonstrating biological plausibility is not part of epidemiological methods. This, however, does not mean that epidemiologists can forget about it. Epidemiologists should understand the biology of the diseases they are studying, explain their hypotheses in biological terms, and propose and promote (sometimes even conduct) biological research to test hypotheses (2016) [33].<sup>20</sup>

— Ultimately, biological processes regulate all pathologies and adverse health effects, without exception. This also applies to social and physiological processes, so that the harmful effects of economic losses must ultimately be carried out through biology (2016) [33].<sup>21</sup>

— Data integration refers to the combination of data, knowledge, or reasoning from different disciplines or from different approaches in order to generate a level of understanding or knowledge that no discipline has achieved alone (2015) [34].<sup>22</sup>

— Today, researchers considering causal inference must pool data from various scientific disciplines (2015) [34].<sup>23</sup>

<sup>16</sup>“...biological plausibility represented fundamental concepts of data integration—the criterion implies that epidemiology and biology must interact” [34] (with reference to [52]).

<sup>17</sup>“All of the relationships and findings should make biological and epidemiological sense” [75].

<sup>18</sup>“The cause—effect relationship should make biological sense” [33].

<sup>19</sup>“Biological plausibility is where biology and epidemiology merge” [46].

<sup>20</sup>“Demonstrating biological plausibility is not part of epidemiological methods. This does not, however, mean epidemiologists can forget about it. Epidemiologists need to understand the biology of the diseases they study, explain their hypotheses in biological terms, and propose and promote (sometimes even lead) biological research to test hypotheses” [33].

<sup>21</sup>“Ultimately, biological processes govern all diseases and adverse health outcomes, without exception. This applies to social and physiological processes alike, so the ill effects of economic deprivation on health must, ultimately, occur through biology” [33].

<sup>22</sup>“Data integration refers to the incorporation of data, knowledge, or reasoning from across multiple disciplines or approaches, with the goal of generating a level of understanding or knowledge that no discipline has achieved alone” [34].

<sup>23</sup>“Today, researchers considering causal inference must integrate data from a variety of scientific disciplines” [34].

— Medical progress develops best when disciplines that focus on subcellular and molecular basic research work in tandem with epidemiology (2003) [31].<sup>24</sup>

In the US judicial system, causality in toxic offenses is sometimes judged on the basis of a wide variety of evidence of varying scientific value—from epidemiology, animal studies, chemical analogies, case reports, regulations, and other secondary sources (2004) [32], although epidemiology is of paramount importance (see [4]).<sup>25</sup>

The Guidelines for Weight of Evidence analysis (WoE; see sources above) used by the International Chemical Safety Program (WHO; toxicology, chemistry, environment, including biota) for assessing impact of agents includes a modified version of Hill’s criteria, with “Biological Plausibility” being listed first (2017) [40]. The fact is that environmental protection agencies (USEPA in particular) form their conclusions sometimes only on the basis of animal experiments [31].

Data integration also applies to radiation disciplines:

— Radiation protection experts, both legislators and practitioners, maintain an understanding of current knowledge in radiobiology and epidemiology to support appropriate decisions (2009) [76].<sup>26</sup>

— When epidemiology reaches its limits, it calls on radiobiology for help (2000) [77].<sup>27</sup>

— “Radiation epidemiology and radiobiology work in tandem for the practice of radiation protection” (2010) [78].

— Integration of fundamental radiobiology and epidemiological research: why and how. Great importance is given to the use of basic radiobiological data in the development of methods for radiation risk assessment (2015) [79].<sup>28</sup>

The meaning of the material presented is clear if we remember that epidemiology itself, similarly to all descriptive/observational disciplines, is almost only established associations [1–5, 78]. As a result, there are

<sup>24</sup>“Medical progress is often best advanced when the sciences that focus on subcellular and molecular basic research work in tandem with the population-oriented science of epidemiology” [31].

<sup>25</sup>“...on a wide variety of evidence of differing scientific value, including epidemiology, animal studies, chemical analogies, case reports, regulatory findings, and other secondary sources” [32].

<sup>26</sup>“...experts in radiation protection, both legislators and practitioners, maintain an understanding of the current state of knowledge in the radiobiology and epidemiology that underpins their subject in order to ensure that the decisions made about regulation are appropriate” [76].

<sup>27</sup>“As epidemiology reaches its limits, it turns to radiobiology for help!” [77].

<sup>28</sup>“Integrating basic radiobiological science and epidemiological studies: why and how.” “A greater role can be established for the use of basic radiobiological data in the development of radiation risk estimates” [79].

facts of the previously cited absurd statistically significant direct or inverse correlations between the consumption of ice cream and mortality from drowning [1], between the export of lemons to the United States and the number of car accidents there [3], between the linear size of the penis in the male population of the country and the level of gross domestic product (GDP) in it [80].<sup>29</sup> The seemingly more real correlations from medicine and epidemiology may only seem so, because the absurd is perhaps just more hidden and intricate.

V.V. Vlasov [24] in his manual writes about "Biological Plausibility" that, in fact, this feature is a variant of the character "explainability of correlation."

Most likely, this definition is the most concise and accurate. It can be used as a guide in everyday life: with all the exceptions, there is little probability that an absurd and inexplicable correlation will be true and a true correlation will be inexplicable.

#### BAYESIAN ANALYSIS (META-ANALYSIS)

A systematic approach based on the integration of data from various disciplines in some works is referred to as "Bayesian analysis" [31] = Confidence Profile Method = Bayesian method [81, 82] and sometimes to "Bayesian meta-analysis" [82–84].<sup>30</sup> It relies on the

<sup>29</sup>Work [80] is not a parody and is not dated April 1 (release date, July). The publication, although called "discussion," is listed as an ordinary one in the database on the website of the Helsinki Center of Economic Research. The study was carried out quite seriously: at the beginning of the article, the author thanks a number of similar, apparently, researchers for "insightful comments" and for providing information about the size of the penis, and the appendices contain all the original data. However, GDP for a sample of 76 countries was associated with racial physiology, which turned out to be an interfering factor (confounder), rather than with the size of the specified organ. Namely, the plot of the dependence of GDP on the length of the penis from [80] shows developed countries of Asia (South Korea, Japan, Singapore, etc.) to the left on top and the undeveloped countries of Africa (Chad, Sudan, etc.) to the right below. The rest of the countries are distributed mostly in the middle. Hence, a high inverse correlation ( $r = -0.447$ ) was revealed. For some reason, this point is not discussed in [80], but there is a hypothesis about the effect of testosterone, the level of which, according to T. Westling, is reflected in the propensity of the population to take risks, including economic ones. This, probably, means that, in Africa, people who are highly inclined to take risks are placing their economy on various dubious "pyramids." In the "Conclusion" of the article [80], the author notes: "Somewhat surprisingly, penile length was a stronger determinant of economic development than the country's political regime type." If this is a prank, then it is very deeply hidden. And it should be added that, according to Hill's criteria, work [80] would never have passed.

<sup>30</sup>"Confidence Profile Method." The only discovered Russian-language source where the indicated English-language title is given is an article in a Belarusian journal from 2005 [83], where there is no translation. However, Bayesian network meta-analysis is often used in medical research in Russia (see, for example, on a Google search).

Weight of Evidence (proofs) rather than on one specific study, following the same principles as Bayesian decision making [31, 81–84]. The method implies inclusion of all available biomedical data (from case reports to animal studies) into one updated and standardized estimate that determines the overall strength of causation (see [31]).

According to authors from the United States [31], the majority of clinicians in their daily practice use Bayesian analysis in conjunction with inductive reasoning, a practice known as differential diagnosis.

#### TELEOANALYSIS: A QUESTIONABLE KNOWLEDGE LINKAGE

Meta- and pooled analyses are used in a wide range of disciplines [85] (and many others), whereas the term "teleoanalysis" is barely known to many people. Three sources for this term can be retrieved from PubMed, and the first of them, by N.J. Wald and J.K. Morris (2003) [86], is foundational. The authors of [86] thank J. Aronson for proposing the term as such (from "teleos," which means "complete"). The literal meaning is a combination of data from *various* types of studies [86]. The difference from meta- and pooled analyses is that data from *different* classes of evidence are combined. The purpose is to implement the linkage (*quantitative*) in terms of the final causation between two such associations, each of which can be proven, but it is difficult to establish the correlation between the first and the final effect of the second for a number of reasons, including ethical concerns (experiments on people).

The following example is given in [86]. Consumption of saturated fat is known to increase the risk of coronary heart disease, but the magnitude of this effect cannot be established experimentally, because long-term trials based on dietary habits are hampered practically. To estimate the overall magnitude of the effect, it is proposed to solve the problem by pooling data using the cholesterol concentration as an intermediate factor in the causal pathway. The effects of (a) reduced fat consumption on the serum cholesterol concentration (which can be done in small experimental studies) and (b) cholesterol concentration on the risk of coronary heart disease (which can be done in epidemiological studies) are investigated. These data are then made quantitatively consistent to assess the effect of dietary fats on the pathology.

The teleoanalytic approach, as far as we can judge, is not very widespread. In addition to [86], we encountered this term in three other papers (PubMed did not identify one of them) [87–89]. Its application (mainly for studying the effect of drugs) was found only in [86], although a similar approach (without the term) can be

found in studies on the assessment of the toxicity of compounds for humans (see, e.g., [34]). However, it should be noted that teleoanalysis differs from simple data integration for revealing biological plausibility in that it is a *quantitative* estimate linking two separate associations, which is questionable. No one can know which quantitative biological patterns at the *in vivo* level will be in a place where it is difficult or impossible to determine them. The conclusion that a certain *quantitative* diet through the concentration of cholesterol will lead to a certain *quantitative* frequency of coronary heart disease, without a direct study of the diet–pathology correlation seems far-fetched and even partly unfalsifiable, because the influence of confounders is likely.

For example, the authors of [90] give a conditional example of a causal chain of events consisting of five units, each of which is characterized by a probability of 90%. In this case, the probability of a common linkage, from the first to the last unit, will be  $0.9^5 = 0.59$  or 59%. If we imagine a 75% probability (which is also quite high), then at the output we will obtain only 24%. It is interesting to know what will remain of the teleoanalysis of the correlation between the diet and the coronary heart disease in this case.

Possibly, this approach is justified only for the “precautionary principle” [1, 51], when at least some data are needed for making decisions.

On the other hand, ideologically, the method of teleoanalysis allows various speculations on the joining of refined and clear fragments of generally unfinished and unclear studies. In addition, the term appears impressive, like meta- and pooled analysis and like “speaking in prose” in Moliere.

#### NON-ABSOLUTION OF CRITERION: WHAT IS BIOLOGICALLY PLAUSIBLY DEPENDSON THE CURRENT BIOLOGICAL KNOWLEDGE

It seems that this Hill’s statement from 1965 [11] (“What is biologically plausible depends upon the biological knowledge of the day”) is often considered as an original thought [28, 29, 49, 87, 91], but it can be found already in works by previous researchers.

In a publication by A.M. Lilienfeld from 1957 [20], the following is stated:

(a) There are historical examples in which a statistical association did not initially conform to current biological concepts. The advancement of knowledge changed concepts that were already consistent with the previously identified associations.

(b) On the contrary, there are examples where an association was interpreted as consistent with existing

biological concepts, but a later interpretation of the association revealed an error.<sup>31</sup>

Similar provisions are expressed in the article by Lilienfeld from 1959 [21] with reference to the previous article of 1957 [20]: the interpretation of any correlation is limited by our biological knowledge, and it may be that an association that does not seem biologically plausible today will turn out to be such when our knowledge expands.<sup>32</sup>

In 1960, P.E. Sartwell [22] noted about his point “the biological reasonableness of an association” that one must be careful with this provision, because judgments made on its basis are supported by imperfect knowledge existing at a particular time.<sup>33</sup>

After Hill, a repetition of this thesis can also be found in serious sources (1991–2018) [15, 30, 43, 50, 67, 92]. Once again, the danger of the situation should be noted: there is a possibility that incorrect and absurd associations declared by someone, as well as data of this kind, can be justified by the fact that the time has not yet come and that science does not yet know.

Historical examples illustrating the two points cited by Lilienfeld (1957, 1959) [20, 21] are of interest:

- The haberdasher J. Graunt investigated the risk factors for plague in London in 1662. All four points of his advice (to avoid contaminated air brought by ships in the port, crowding, and contacts with animals and sick people) turned out to be constructive, despite the absence of any biological theories [48, 93].

- An extreme increase in scrotal cancer in English chimney sweeps was identified by P. Pott in 1776 [11, 15, 48, 93], 150 years before the start of research on chemical carcinogens. However, after the work by Pott, the Danish parliament issued a decree on “prevention”: chimney sweeps were ordered to wash every day [15].

<sup>31</sup>“There are historical instances in which a statistical association did not originally conform to existing biological concepts. As advances in knowledge changed the biological concepts, these new concepts were found to be consistent with the previously observed association. Conversely, there have been instances in which the statistical association was interpreted as being consistent with existing biological concepts, but later the interpretation of the association was found to have been erroneous” [20].

<sup>32</sup>“It is unnecessary to point out that our interpretation of any relationship is limited by our biological knowledge, and it may well be that an association which at first does not appear to be biologically plausible will turn out to be so when our knowledge has been extended [4]. In fact, the finding of a biologically implausible association may be the first lead to this extension of knowledge. However, all that can be done is to indicate that the association should be investigated further” [21] with reference to [20].

<sup>33</sup>“The Biological Reasonableness of the Association—this point cannot and should not be left out of consideration, but we must be wary of it, for judgments made on this basis are hemmed in by the imperfect knowledge existing at any time” [22].



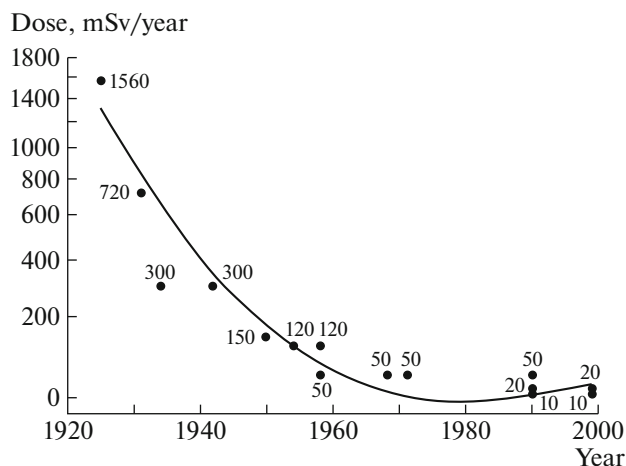
- In 1842, D. Rigoni-Stern, having analyzed the statistics for Verona, discovered that mortality from cervical cancer was characteristic of married rather than single women, indicating the influence of sexual or reproductive activity. The fact was forgotten until the early 1950s, when the rarity of this pathology among Catholic nuns was noted. In the 1960s, a search for a disease-causing factor associated with sexual behavior was carried out, and in the 1970s the herpes virus (HSV II) was proposed as such a factor, which at that time was regarded as a biologically plausible hypothesis. However, subsequent epidemiological studies have not confirmed this correlation. In the 1980s, with the development of the DNA hybridization method, the true factor was discovered. It turned out to be human papillomavirus (HPV), although of specific genetic subtypes [70]. The author of the review [70] noted that it is difficult to imagine how the evidence would have been obtained without a fruitful intersection of epidemiological and biological approaches.

- In 1848–1849, W. Farr found an inverse association between living height above sea level and cholera mortality in London [94]. The phenomenon was consistent with the theory of “miasmas,” which prevailed at that time, and was interpreted as plausible. Later it was found that, according to the microbial theory, the situation is the same: the height of water sources was inversely correlated to the etiological factor (in publications on the topic, an example is given in [20]).

- In 1849–1854, J. Snow identified a correlation between drinking water pollution and cholera incidence in London [95]. Snow’s findings did not appear biologically plausible in light of the miasma theory. However, when the microbial theory became generally accepted, the correlation between water pollution and cholera became scientifically substantiated (in publications on the topic, an example is given in [20] and later repeated in [11, 68, 96, 97]).

- In 1861, D.W. Cheever (United States), warning about the danger of “meaningless correlations,” pointed out that it would be ridiculous to attribute typhus, which someone caught while spending the night on an emigrant ship, to parasites on the bodies of sick people. In his opinion, this was just a coincidence [98] (in publications on the topic, the example was originally given by Sartwell in 1960 [22] and later repeated by Hill [11] and then by Rothman and Greenland [43, 45, 99]).

- At the dawn of AIDS research, its causes were associated with the use of amyl nitrites (“poppers,” inhaling of which increases, say, temporary temperament) by homosexuals [100].



**Fig. 1.** The evolution of radiation protection standards (RPSs). Plotted (Statistica, ver. 10) according to U.S. Department of Energy (DOE-1995) [102], in abbreviation. For 1958, 1990, and 1999, RPSs are presented for nuclear workers and the public; for 1990, RPSs are presented from BEIR-V (50 mSv) and ICRP (20 and 10 mSv), supplemented by RPS-99 (Russia). The initial doses in [101] were presented in rem/year (rem is roentgen equivalent man; 1 rem = 0.01 Sv [103]). The values of dose standards are shown on the plot.

These are the most popular examples (apart from the second and the last one) regarding the temporal conjuncture of plausibility, which are given in the textbooks on epidemiology and reviews. However, we can recall more cases from the history of medicine, at least with the therapy of all diseases by bloodletting or with the drinking of a living culture of *Vibrio cholerae* by M. von Pettenkofer without consequences, which “confirmed” the hygienic theory of “miasmas” [101].

The concepts of biological plausibility have evolved greatly in both radiation epidemiology and radiobiology. Figure 1 shows the development of global radiation safety standards (RSS) based on data from [102].

The decrease in the permissible dose by two orders of magnitude in less than 70 years fully illustrates the evolution of ideas about the biological plausibility of the effect of radiation. At first, the concepts of “tolerance doses” (until 1942) and then the “maximum permissible exposure” (until 1950), meaning something “harmless,” were used. In 1950, the concept changed to the harm of irradiation at any dose, followed by a reduction in the dose standard as much as it is “as low as reasonably achievable” [102].

While the concept of the threshold radiation harmlessness prevailed (note that that “threshold” was very high), the ideas about its effects were, in the modern view, anecdotal in nature. Radiologists during the First World War scanned the patient’s chest standing

directly behind him and holding the X-ray film in their hands [104].

“Healthy” radium-containing water “Radithor” (2  $\mu\text{Ci}$  in 16.5 mL per bottle; half  $^{226}\text{Ra}$  and  $^{228}\text{Ra}$  [105, 106]), which in 1932 the famous celebrity E.M. Byers (entrepreneur and golfer) [105–108]<sup>34</sup>, as well as rectal suppositories with radium (1930), which “charge like a battery” [114, 115] (recall the advertisement for the “Energizer” bunny), are not the most egregious examples.<sup>35</sup> Popular in France in the 1930s were face creams and other perfumes “enriched in thorium and radium,” manufactured under the Tho-Radia brand, as well as the similar Doramad toothpaste in Germany [105, 116], surprising, though not overly so, after all that happened in those days. Radium was then added everywhere: in chocolates, in eye drops, in Christmas tree decorations, and in children’s books with luminous letters. Peculiar siphons to enrich drinking water with radium were widely used. Many such examples are given in the manual on radiation ecology by G.N. Belozersky published in 2019 [116], and photos of all these wonders can be found in Russian and English-language Internet blogs.

However, most impressive is the relatively recent (1950–1951) children’s construction game “Gilbert

U-238 Atomic Energy Lab” (there are also other similar games), with teases like “Build yourself a nuclear power plant, buddy!” (“Showing you the real atomic decay of radioactive material!” [117]).<sup>36</sup> The box with the game, in addition to the electroscope, etc., included containers with four types of uranium ores, as well as with  $^{210}\text{Pb}$ ,  $^{106}\text{Ru}$ ,  $^{65}\text{Zn}$ , and  $^{210}\text{Po}$  [117].

The evolution of ideas about hereditary genetic changes in humans after radiation exposure (i.e., about defects, pathologies, and anomalies in unirradiated children of irradiated parents) is also very illustrative. Such changes were found first in *Drosophila* (1927) and then in mice and other animals (1950s) [118, 119]. The world froze in anticipation of the mass of radiation mutants in people as the results of nuclear weapons tests, radioactive pollution of the environment from nuclear energy, the progressive use of radiation in medicine, etc. All this was reflected in the mass media, cinema, and science fiction literature of the 1960s–1970s (and later in computer games), where horrible mutants with incredible abilities appeared [119].

As time passed, mutants were not born even in the victims of Hiroshima and Nagasaki, nor in tens of thousands of survivors after radiotherapy. Finally, it was concluded that the hereditary genetic effects in irradiated people cannot be detected in real studies, because even the theoretical increase in the natural mutational background is so small at any plausible doses that do not eliminate fertility. If at present someone detects radiation mutants in humans, this does not meet the plausibility criterion (more precisely, coherence with the available epidemiological data), as well as the provisions of the UNSCEAR, ICRP, and BEIR [8, 118–122].

In fact, according to recent data of Russian authors, at the level of microsatellite polymorphisms, gene exons, and mtDNA heteroplasmy, such a possibility was apparently shown for irradiation with high doses (2–3 Gy) shortly before conception (workers of the Mayak PA) [123, 124].<sup>37</sup> However, in other pilot studies by the groups headed by A.V. Rubanovich (Kuzmina N.S. et al., 2014; 2016 [125, 126]), no transgenerational transmission of even epigenetic changes (hypermethylation of CpG islands in the promoters of a number of genes) was detected in the descendants of the liquidators of the Chernobyl accident and workers in the nuclear industry, which fits into the big picture. But, for the parental contingent itself, including the

<sup>34</sup>Having consumed about 1400 bottles of “Radithor” (~2.8 mCi) over three years, together with his girlfriend, E.M. Byers lost his jaws and partly his skull and died a terrible death. Six months later, his girlfriend died from the same cause. Today, the remains of this victim of trust in “current scientific knowledge” are buried in a cemetery in Pittsburgh, in a lead coffin (the half-life of  $^{226}\text{Ra}$  is 1600 years) [106, 107]. However, the “current scientific knowledge” (“biological plausibility”) even then testified to something somewhat different; namely, the fact that exposure to  $^{226}\text{Ra}$  can have destructive effects on blood cells [109] and rodents [110] was shown as early as the 1920s. For example, in the work by K. Sugiura and G. Failla in 1922 [110], it was demonstrated that a relatively low exposure to  $^{226}\text{Ra}$  (2.4  $\mu\text{Ci}/\text{h}$ ) stimulated the growth of mice (with concomitant permanent sterilization of females, which is clearly undesirable), whereas irradiation at 21.9  $\mu\text{Ci}/\text{h}$  had the opposite effect, and at 36.5  $\mu\text{Ci}/\text{h}$  the animals died quickly. Even earlier, in 1913, W.C. Lazarus-Barlow discovered that the final depot for  $^{226}\text{Ra}$  is the bone tissue [111] (cited from [107]). Then, the same author in 1924 assumed the role of  $^{226}\text{Ra}$  as a cause of cancer [112]. Thus, there were sufficient scientific reasons not to drink radium water, especially in such quantities, by the mid-1920s. However, “Radithor” was prescribed to Byers in 1927 by a physician when the golfer broke his arm as a result of fall in a train “to aid healing” [107]. Byers liked the elixir of distilled water and  $^{226}\text{Ra}$  +  $^{228}\text{Ra}$  so much and felt so “invigorated” that he got his dose without a doctor [106, 107]. Presumably, it had a potency stimulation effect, because this person was known for his adventures [106, 113]. At that time, “Radithor” was taken by many important people [106], though probably not in such quantities, and therefore other information about the consequences is not known to us.

<sup>35</sup>“Vita Radium Suppositories Popular during the early 20th century, Vita Radium suppositories contained real radium, the radiation of which was said to charge up the body like ‘charging up an electric battery.’ The suppositories were claimed to cure fatigue, impotence, low sex drive, and hemorrhoids”) [114].

<sup>36</sup>“...shows you actual Atomic disintegration of radioactive materials!” [117].

<sup>37</sup>There are no data yet that these subtle changes can be realized in tangible disorders, anomalies, and pathologies, which have not been registered in tens of thousands of offspring in various cohorts over more than 60 years [8, 118–122].

workers at the Mayak PA, the above-mentioned epigenetic changes depending on radiation exposure and age were clearly identified [125–127].<sup>38</sup>

Thus, the influence of temporal conjuncture on the biological plausibility is evident in many disciplines. Nevertheless, according to Daubert's rule [4, 31, 32], expert testimony in US courts should be judged on the basis of current scientific knowledge only, and not on the possibility that additional information may become available in the future.<sup>39</sup> No guesswork as to what new discoveries might reveal is allowed: the courtroom is not a place for scientific guesswork, not even as inspiration. The law lags behind science, but does not lead it [32].<sup>40</sup>

#### NON-ABSOLUTION OF THE CRITERION: REAL BUT UNPLAUSIBLE ASSOCIATIONS AND PLAUSIBLE BUT NOT REAL ASSOCIATIONS

Earlier, we gave examples of the effects of confounders [1, 4], chances, and reverse causality [3] in the formation of associations. Correlations such as ice cream consumption—drowning frequency, the correlation between alcoholism and lung cancer, between smoking and liver cirrhosis, between the use of amyl nitrites and the incidence of AIDS, between smoking cessation and an increase in pneumonia and lung cancer, as well as other correlations considered are all examples of the first part of the heading—the associations seem to be real, but their biological (and sometimes logical) plausibility is absent.

More similar facts can be added.

(1) There was a deplorable situation with thalidomide, which was not a teratogen for most experimental animals, but in 40% of cases it was teratogen when taken by pregnant women, causing severe pathologies such as underdevelopment of the limbs, etc. (report of 1961) [93, 128].

(2) The second tragedy (the authors of [129] call it "intellectual pathology" of physicians") is associated with the rather long-term use (1942–1954) of oxygen-

ation for premature infants (high-dose oxygen therapy). Although the biological basis and a plausible mechanism were well understood, nobody could have foreseen the consequences—retrolental fibroplasia (overgrowth of connective tissue under the lens) leading to blindness [35, 129]. The procedure has blinded approximately 10000 children. Evidence of causality was difficult to obtain, because the defect was initially assumed to be of neonatal origin. Associations were also found with fetal abnormalities, multiple gestations, maternal infections, light, vitamin and iron supply, vitamin E deficiency, and hypoadrenalism. In addition, in animal experiments, no data confirming the corresponding effect of both hyperoxia and hypoxia were obtained. However, some time later, both a cohort study showed that the defect developed in children with initially normal eyes and a counterfactual experiment involving the cancellation of exposure (for the concept, see [1, 11], etc.), which led to a sharp decrease in the iatrogenic epidemic, demonstrated the reality of the correlation (however, with the cessation of oxygenation, an increase in the mortality of premature infants from respiratory distress syndrome was observed) [129].<sup>41</sup>

(3) The recommendation by Dr. B.M. Spock to put infants to sleep prone to reduce the risk of sudden infant death syndrome is based on the plausible explanation that it prevents suffocation with vomiting. However, it turned out that mortality in the supine position was two times lower than in other positions [33, 90].

In a number of cases, the results of randomized controlled trials (RCTs) fundamentally contradicted the data of numerous epidemiological studies, which significantly and systematically reveal certain associations (by the way, biologically plausible). The publication by C. Lacchetti et al. in 2008 [130], which is included in the American Medical Association guide-

<sup>38</sup>The authors of [125, 126] indicate that, since the age of the offspring did not exceed 40 years and the epigenetic changes depended on age, there is a possibility that the frequency of transgenerational transmitted changes may be accelerated in the future.

<sup>39</sup>The basics of the so-called "Daubert ruling" ("Daubert standard") were laid in 1993, when the parents of children born with serious defects, J. Daubert and E. Schuller, filed a lawsuit against the pharmacological company, claiming that the use of the drug Bendectin by pregnant women led to the defects [4, 31, 32]. The court considered all pertinent evidence from the biomedical disciplines.

<sup>40</sup>"The courtroom is not the place for scientific guesswork, even of the inspired sort. Law lags science; it does not lead it" [32].

<sup>41</sup>It is pertinent to add here that, as indicated in the manual by Rothman from 2008 [43], when determining the risk of mortality from an interventionist factor, the entire range of possible causes (counterfactual [1]) should be assessed. In [43], an example is given with oral contraceptives, which, in principle, can increase the risk of death within ten years (for example, from myocardial infarction or stroke). However, it is necessary to consider immediately the alternative to taking contraceptives, which may be pregnancy. And pregnancy and childbirth also have their own risk of death, which may be higher than that from contraceptives. In this regard, in practice the latter will turn out to be a preventive, but not causal, factor of mortality. This holds only if the alternative is another contraceptive, without side effects (but not childbirth), then the first contraceptive drugs, with shortcomings, can be regarded as the cause of death. Equally, the fact that the abolition of oxygenation in premature infants, having eliminated the risk of blinding, simultaneously led to an increase in mortality from respiratory distress syndrome [129], suggests that, in this case, too, the risks should be compared and weighed before making any judgment.

lines on the use of medical literature in evidence-based medicine, gives 22 such examples.

In principle, RCTs are regarded as the “gold standard” of evidence [58, 67, 93, 96]. Although they have many limitations [131–133] (and many others), so that the “gold standard” of RCTs is directly denied in a number of works [134–136], the official doctrine assigns this purely experimental approach the highest rank among the methods of evidencing causality for humans (see manuals for the past two decades [58, 93, 96]). Below are the most well-known examples of the discrepancy between the results of RCTs and observational studies.

(4) Potassium fluoride reduced spinal fractures in an observational study in 1989, but not in the subsequent RCT in 1990 (see [131]).

(5) Epidemiological and laboratory data led to the conclusion that  $\beta$ -carotene reduces the risk of cancer (article in “Nature” in 1981 [137]). However, two RCTs in 1990s demonstrated that  $\beta$ -carotene increased lung cancer mortality in smokers, although this contradicted the entire body of epidemiological and biological evidence (see [51]).

(6) Vitamin E in observational studies in 1994 reduced the risk of coronary heart disease, but an RCT in 1994 did not confirm this effect (see [131]).

(7) The most well-known example of the inconsistency between the results of RCTs and observational epidemiological studies is hormone replacement therapy (HRT) in menopause. This example, mainly with an emphasis on the absoluteness of RCT data (with few exceptions [96]), is discussed in textbooks and guidelines on epidemiology and evidence-based medicine [24, 35, 93, 96, 131, 138]. Early observational studies demonstrated the beneficial effects of estrogen therapy, including a reduction in the incidence of coronary and cardiovascular pathologies, as well as in overall mortality (six case-control studies, three cross-sectional studies, and 15 prospective studies; see the review by M.J. Stampfer and G.A. Colditz from 1991 [139]). It should have been alarming, however, that the incidence of mortality from homicide and suicide in women with HRT was also reduced: this implausible dependence is a negative control for the presence of confounders (for example, lifestyle, habits, and social status) [142]. Indeed, the very first RCT in 1998, as well as the subsequent RCTs (2002–2012; see [138, 143]), did not confirm the effects. Moreover, an increased risk of coronary diseases, strokes, thrombosis, pulmonary embolism, and breast cancer was found [35, 138, 143, 144], so that one of the RCTs had to be ceased before the planned date [144]. All this was reflected in the media, sometimes in a distorted form (according to [143], the relative risk of breast cancer (1.24) was transformed into

a 24-fold increase in the risk of breast cancer after HRT). Nevertheless, the positive effects of HRTs prevail over the negative ones for women aged 50–60 years, whereas at an older age mainly adverse consequences become significant (see [96, 143]).

There are other, no less spectacular facts from other areas.

(8) For example, the lack of a correlation between psychoemotional stress (including severe stress, such as the death of children and loved ones [145]), including depression, and an increase in the risk of various types of cancer [145–149] seems strange. Associations between psychological stress and malignant breast tumors, found in some studies [145, 150], do not meet the criterion “consistency of association” (not found in a large Danish study [151]; see also [145, 147, 150]). However, the biological plausibility here reaches a high degree of significance: depression of the immune system, hormonal (neuroendocrine) shifts [93, 146, 152, 153], and oxidative stress associated with psychoemotional stress, which leads to DNA damage and cytogenetic disorders [121, 150, 152–154]. Moreover, experiments on animals have repeatedly demonstrated that psychoemotional stress in various models actually increases the level of DNA damage (see [121, 150, 154]).

However, epidemiological studies have failed to demonstrate this association [145–149]. Nevertheless, it is indicated that stress can act indirectly, reinforcing bad habits and unfavorable lifestyle, which may cause cancer [146, 148]. Theoretically, stress can also promote tumor growth and progression through decreased immunity and changes in hormonal status [148, 152, 155, 156]. However, stress correction in cancer patients did not change in their survival [148, 157].

It can be seen that all of the listed biological plausibility [93, 121, 146, 150, 152–154] does not necessarily indicate epidemiologically real consequences [145–149]. On the website of the National Cancer Institute of the United States, the answer to the question as to whether psychological stress can cause cancer states that the evidence is weak and inconsistent [148]. On another official website of Cancer Research UK, the world’s largest cancer research charity from England, it is also written that the majority of studies did not show an increase in cancer risk due to stress [149]. Finally, in the IARC document (IARC 2014; updated 2018), it is noted (probably with reference to [147]) that meta-analysis of data for 12 European cohorts did not reveal an increase in cancer incidence (including major cancer types) under various types of stress at work [158].

(9) The mechanism of the anesthetic effect of acupuncture is unclear. Many researchers consider it biologically implausible that the insertion of needles into

the body and their rotation there can relieve pain and, therefore, do not believe in the effectiveness of acupuncture [58].

(10) Similarly, the effect of homeopathy is regarded as not having a biological mechanism [87] (except for the weak and long exaggerated [159] placebo effect [87]).

#### NON-ABSOLUTION OF THE CRITERION: A BIOLOGICALLY PLAUSIBLE MECHANISM CAN EASILY BE FOUND FOR EXPLANATION OF EACH ASSOCIATION

This is a statement from the editorial by G. Davey Smith in 2002 [160], which is a response to a certain "standard argument" according to which hypotheses based on a good scientific understanding of pathogenesis can hardly be false. Here, it is appropriate to draw an analogy with the same statement for the criterion "Analogy," about which K. Rothman (including with various coauthors) said that "analogies abound" [74], since they are associated with the imagination of researchers [42, 43, 45, 99].

Such statements, in our opinion, hardly make sense, because they contradict the above thesis reproduced in many epidemiological sources: proof of causality is ultimately a matter of judgment [16, 37, 38, 67].

#### CRITICISM OF THE CRITERION: LACK OF SCIENTIFIC STANDARDIZATION, SUBJECTIVISM, DOGMATISM, STATEMENT THAT THE COMBINATION OF EVIDENCE FROM VARIOUS DISCIPLINES CANNOT BE AN INDEPENDENT DISCIPLINE, AND OTHER DISAGREEMENTS

We are aware of individual authors who have criticized the inductive approach to proving causality using guiding principles or criteria. Actually, in all hundreds of sources of different levels, no others were found. Therefore, most likely, there are no other such authors, at least among those that are known.

(1) Chronologically first is the study of 1979 [161] performed by a wide-profile critic, creator of the term "protopathic bias" (see [3]), one of the "fathers" of clinical epidemiology [162], A.R. Feinstein (1925–2001, United States). In the previous section, we mentioned the work by Jacobson and Feinstein from 1992 [129], in which the "intellectual pathology of physicians" who biblically blinded 10 000 infants by oxygenation, was considered. We also know the publication by Feinstein about meta-analysis, where this approach (but not pooled analysis) is called "statistical alchemy of the 21st century," because it is "the idea of getting something for nothing, while simultaneously ignoring established scientific principles" [163]. Feinstein also did not pass over Hill's criteria [161]. As

applied to "Biological Plausibility" (in the form of "Coherence"), in [161] he noted that this demand provides for biological logic. The latter, however, is based on paradigmatic appropriateness rather than on rigorous evidence. The author concluded that there are no scientific considerations or scientific standards when applying the criteria of causality, because scientific principles are mentioned neither among the statistical nor among the logical components of the complex [161].<sup>42</sup>

According to Feinstein, the inclusion of some experimental "standards," similar to those contained in Henle–Koch's postulates of infectious diseases, will lead to the appearance of scientificity [161]. However, as can be seen in Report 2 [2], Henle–Koch's postulates are also not amenable to standardization.

In general, Feinstein's criticism of the "Biological Plausibility" (and "Coherence") criterion can be formulated as follows: since the criterion is based only on biological logic, it requires acceptance of the initial paradigms of appropriateness and requires judgment, but not rigorous proof. Therefore, the criterion really does not work. This conclusion by Feinstein, however, abolishes the acceptability in practice of the evidence-based methodologies used by environmental agencies [7, 15, 38, 50, 164] and IARC [57, 165, 166], as well as by other organizations that use weight-of-evidence approaches (WoE; see sources above) and Bayesian analysis (meta-analysis) discussed here earlier.

(2) The second is Professor of Epidemiology K.J. Rothman (born in 1945) [41, 42, 44, 74, 167] and his frequent coauthor, no less famous researcher of causality, Professor of Biostatistics and Epidemiology S. Greenland (born in 1951) [43, 45, 65, 99, 168], both from the United States, whose works were mentioned by us many times. We hope to consider their platform in more detail in Report 4. Here, we will only mention briefly that these authors completely deny the inductive approach and, judging by all the signs, the probabilistic causality, reducing everything to a finite multifactorial determinism [41–43, 45, 74, 99] (we have already discussed this in part in [2]). For Rothman, such views are found at least since the first edition of his "Modern Epidemiology" in 1986 [41] (information is taken from the manual [169]) and are then reproduced in the second edition of 1998 [42] (guiding principles are burdened with reservations and exceptions; cited by [170]) and in the third edition of 2008 [43].<sup>43</sup> We had only the latter at hand.

<sup>42</sup>Requirements for the types and quality of studies that will be used and for their design include a rigorous set of methods for eliminating chances, bias, and confounders and for the quality of basic data [161], that is, some quite general requirements.

<sup>43</sup>"The Hill guidelines... are saddled with reservations and exceptions (Rothman 1998, p. 27)" [42]; cited in [170].

Rothman and Greenland for each criterion, including the need for the association itself, found a certain exception, sometimes speculative.<sup>44</sup> As a result, it turned out that no one criterion (except for “Temporality,” which was recognized) is an even weak “point” that allows any judgments [41–43, 45, 74, 99]. In other words, Rothman and Greenland always offered a rebuttal to eight of Hill’s nine propositions, somewhat similar to the argument of a character in Soviet cinema “What if he was carrying rifle cartridges?” (or, even worse, the character by V.M. Shukshin who “cut off”; both sentence from Russian film and book respectively).

However, the authority of Rothman and Greenland in a kind of “theoretical epidemiology” (the current system of epidemiology is largely determined by Rothman’s thinking; manual of 2016 [169]), that is, in epidemiology materials reflected not in the methodological documents of organizations that make decisions and are responsible for them but in textbooks or reviews, is quite large.<sup>45</sup> Rothman is often referred to in the latter as having buried causal criteria. Based only on manuals and reviews, sometimes one can make the erroneous conclusion that today no one uses any causal criteria. We plan to consider in detail the issue of the breadth of application of Hill’s criteria and, in our opinion, their logical inseparability from any evidence in descriptive disciplines in Report 4.

As for “Biological Plausibility,” the authors mentioned call it an “important problem.” However, in their opinion, this criterion is both biased and non-absolute, since it is often based only on previous beliefs rather than on logic or data. According to Rothman and Greenland, attempts to quantify on a scale from 0 to 1 the probability of what is based on previous beliefs and what is based on new hypotheses using a Bayesian approach demonstrate dogmatism or adherence to the current public fashion.

<sup>44</sup>It has already been considered in [1, 3]; however, due to the nonordinary thinking of those who proposed such an exception, it is advisable to recall it here as well. The work by Greenland and Robins from 1986 [168], cited from [60], gives a “bivariant counterfactual” example. It is assumed that half of the individuals in the population are sensitive to some exposure and may die from it, and the other half may die precisely because of the absence of such an exposure (imagine a population where, say, half are heavy drug addicts, for which the required dose of the potion is fatal to the average person). If exposure is distributed randomly across the population, then the expected mean causal effect is zero: there is no association between exposure and mortality in an infinitely large group. However, the observed result for each individual will be determined causally by the fact of exposure or nonexposure [60, 168]. Thus, even an association is not necessary for a causal relationship. It remains to wait for kind of a Pelevin–Orwellian model, which will show that causality itself is not needed to establish causality. However, this is, rather, from the field of lawless justice.

<sup>45</sup>“While the Hill criteria give some useful heuristics for determination of causality, the current system of epidemiology is very much determined by the thinking of Rothman” [169].

This leads to bias in evaluating the hypothesis [45, 99]. Neither Bayesian [45, 99] nor any other [43] approach can turn plausibility into an objective causal criterion [43, 45, 99].<sup>46</sup> In assessing a new hypothesis, this criterion can be used only in a negative sense (in order to indicate the difficulty of its application) [45, 99].<sup>47</sup>

In the last publication by Rothman on this topic known to us (2012), the characteristic of the criterion of interest is very brief: “very subjective” [74].

(3) The next in time (after the first edition of “Modern Epidemiology” by Rothman in 1986 [41]) is the publication by Professor of Theoretical Medicine B.G. Charlton in 1996 [171] (England). In 2019, according to data from the Internet, Charlton was still teaching. Judging by PubMed (last publication in 2012), he is also a wide-profile critic with rather unusual works: “The Zombie Science of Evidence-Based Medicine...,” 2009 [172]; “Are you an Honest Scientist?,” 2009 [173]; “The Cancer of Bureaucracy: How It Eill Destroy Science, Medicine ...,” 2010 [174], etc. For example, the following statements are impressive: “it is obvious that Evidence-Based Medicine was, from its very inception, a zombie science: reanimated from the corpse of clinical epidemiology” [172].

Regarding the significance of the criteria of causality, Charlton, similarly to the above-mentioned researchers, takes quite extreme positions of denial (like the firemaster of the Russian classics, who was so right-wing that he did not even know which party he belonged to).

Speaking about multidisciplinary and interdisciplinary approaches in epidemiology, Charlton notes their mosaic nature. Each piece of the mosaic consisting of specific data can be assessed for scientific validity, but the method by which these elements are connected (glued together) is not scientific, because, according to [175, 176], a combination of evidence from different disciplines cannot be an independent scientific discipline. If gaps in evidence for one disci-

<sup>46</sup>“Plausibility refers to the biological plausibility of the hypothesis, an important concern but one that is far from objective or absolute.... Such is the problem with plausibility: it is too often not based on logic or data, but only on prior beliefs.” “The Bayesian approach to inference attempts to deal with this problem by requiring that one quantify, on a probability (0 to 1) scale, the certainty that one has in prior beliefs, as well as in new hypotheses. This quantification displays the dogmatism or openmindedness of the analyst in a public fashion, with certainty values near 1 or 0 betraying a strong commitment of the analyst for or against a hypothesis. It can also provide a means of testing those quantified beliefs against new evidence. Nevertheless, the Bayesian approach cannot transform plausibility into an objective causal criterion” [99]. “Nevertheless, no approach can transform plausibility into an objective causal criterion” [43].

<sup>47</sup>“This is not to say that biological knowledge should be discounted when evaluating a new hypothesis, but only to point out the difficulty in applying that knowledge” [99].

pline are replaced or circumvented using data from another discipline, and vice versa [177], then incompatibility is combined. Indeed, since data obtained using a number of "incommensurable" (or qualitatively different) approaches, methodologies, and methods of evidence are integrated, epidemiology should rely on judgment, on "common sense," to a greater extent than other sciences [171].

Charlton compared evidence in epidemiology with a "network" of related data from various disciplines. Insufficiency or incorrectness in any cell does not break and does not eliminate the entire network, but only weakens it, because it is a *network*. That is, epidemiological hypotheses are very weakly falsifiable (according to the provisions by K. Popper [1, 37]): conflicting results can only change the equilibrium probability of multifactorial causation. This explains the long life and stability of other epidemiological hypotheses even in the face of conflicting facts (see the examples above in the previous section). On the contrary, "true" scientific hypotheses are built on a *chain* of evidence and separate evidential links and, therefore, the elimination of any link as incorrect immediately completely eliminates the hypothesis itself [171].

Among the sources cited by Charlton on the unscientific integration of disciplines, publication [175] is a 1987 monograph entitled "Molecules and Mind" and work [176] published in 1990 belongs to Charlton himself; it is devoted to criticism of "biological psychiatry."

In general, Charlton's reasoning is not devoid of sense when it is not about epidemiology, but, for example, about psychology or, say, about retrospective disciplines unverifiable in practice. This also applies to the cases when pretensions for *quantitative* connections between fragments of epidemiological evidence are criticized, as declared for teleoanalysis (see above). However, in the case of an approach predominantly in a qualitative sense, which provides, as said, "the Weight of Evidence," such absolutist reasoning renders senseless many practical steps of international organizations aimed at preserving health and the environment.

The considered fairly old publication by Charlton from 1996 [171] is cited by a number of authors dealing with the philosophy of epidemiological causality. It seems appropriate (see why below) to provide a summary of such sources with an indication of the meaning of the citation:

- D.L. Weed, 1997 [59]. Information about the authority Weed has been cited by us several times earlier [2, 3] (see also note 3). Charlton, on the other hand, stated that the use of causal criteria diminishes the validity of causal inferences (in fact, the statements

in the original [171] are not so unambiguous; see the original quote in [2]).<sup>48</sup>

- M. Parascandola and Weed, 2001 [178]. The first author is also very authoritative for his publications on the topic. "Charlton also states that basic science is built on the concept of necessary causes and that epidemiology, in order to be scientific, must follow this model."<sup>49</sup>

- J.C. Arnett, 2006 [179]. Author of the Competitive Enterprise Institute (CEI), discussion of USEPA approaches. OMB [The Office of Management and Budget] just aggregates or compiles estimates from various agencies, but never reviews or endorses all the "different methodologies" used and does not provide valuable and successful estimates (Charlton, 1996).<sup>50</sup>

- A.Z. Bluming, 2009 [180]. Presentation from the University of California at the Symposium on HRT, as well as an accompanying publication in the journal. An excerpt from Charlton [171] about the "true" hypothesis from a chain of causes and about the epidemiological hypothesis, which is a "network" of connected data from many disciplines is presented (see above).<sup>51</sup>

- A. Morabia, 2013 [181]. Again, a highly cited author, see [1, 2]. Cited from Charlton [171] in light of the discussion of criteria after 1965.

- US manual on statistics and causality published in 2016 [169]. So far Rothman (1986) did not criticize all assumptions [about criteria] as inconclusive (see also Charlton in 1996).<sup>52</sup>

- P.J.H.L. Peeters, 2016 [182]. A voluminous (246 pp.) report of a study on European programs. A total of 13 citations of Charlton 1996 [171], most on the topic and without discussion. For example, the

<sup>48</sup>"Charlton, on the other hand, argued that using causal criteria diminishes the validity of causal inferences" [59].

<sup>49</sup>"Charlton similarly claims that the basic sciences are built on the concept of necessary causes and that epidemiology, in order to be scientific, should follow that model" [178].

<sup>50</sup>In 1997 Charlton stated that the OMB [The Office of Management and Budget] simply aggregated or compiled various agency estimates; it neither reviewed nor endorsed all the "varied methodologies" agencies use to derive cost and benefit estimates [179].

<sup>51</sup>"Disproving even a single item of evidence can, in principle at least, overthrow a whole scientific hypothesis, because a causal chain is only as strong as its weakest link. By contrast, epidemiological hypotheses are supported by a network of linked evidence from numerous disciplines, and cutting any single strand may weaken a net, but does not break it. Contradictory findings cannot do more than alter the balance of probability of multifactorial epidemiological causation. This explains the long life, resilience, and apparent irrefutability of such epidemiological hypotheses in the face of powerful items of apparently contradictory evidence" [180].

<sup>52</sup>"...until Rothman (1986) not only criticized all of the assumptions as nonconclusive (see also Charlton, 1996)" [169].

way by which epidemiological evidence is weighed was criticized by Charlton (1996, p. 106).<sup>53</sup>

We had seven more sources (1997–2011) where the provisions from [171] are mentioned/cited, but these works are either older or appear less sound.

Thus, there are 14 sources (half of them are definitely sound), in which the sharply negative provisions by Charlton from 1996 [171] are considered without any criticism. It turns out again that, given the necessary conjuncture, an interested person with the help of authoritative sources, having added the publications by Rothman and Greenland, can not only refute the criterion “Biological Plausibility” but also the majority of Hill’s “logical” criteria, as well as put into doubt the adequacy of the approaches and conclusions of USEPA, IARC, etc., or other studies that synthesize data from different disciplines or are at the intersection of them.

(4) The last of the critics known to us are Professors K.J. Goodman and C.V. Phillips from Canada, who are frequent authors of articles on causality in various kinds of encyclopedias and textbooks [30, 183]. It should be noted that Phillips has ties to tobacco companies [184, 185].

A 2005 article by these researchers said that, with the exception of Temporality, none of the criteria is necessary or sufficient. And it is not clear how to quantify the degree of importance of each criterion, let alone to generalize such an approach to the judgment of causality [183]. Regarding the “Biological Plausibility” criterion, it can be noted that Goodman (with another co-author) considers it together with “Coherence” and “Analogy” [30].

An important statement is the following: there appears to be no empirical assessment to date of the validity or usefulness of causal criteria (e.g., retrospective studies of how the use of criteria improves analytical conclusions). In short, the value of the set of criteria for causal inference is strictly limited and has not been tested [183].<sup>54</sup>

Similar provisions can be found in other authors, for example, in G.B. Gori from 2004 [186]: deprived of both quantitative and qualitative benchmarks, these criteria remained subjective and unrelated to independent experimental verification.<sup>55</sup>

<sup>53</sup>“This manner in which epidemiological evidence is weighted has led to the criticism that, in the words of Charlton (1996, p. 106)...” [182].

<sup>54</sup>“Furthermore, there appear to be no empirical assessments to date of the validity or usefulness of causal criteria (e.g., retrospective studies of whether appealing to criteria improves the conclusions of an analysis). In short, the value of checklists of criteria for causal inference is severely limited and has not been tested” [183].

<sup>55</sup>“Bereft of quantitative and qualitative benchmarks, these criteria have remained judgmental and are not linked to independent experimental verification” [186].

This, so to speak, is the quintessence of the entire criticism of the set of criteria: it is declared that the success of their application *has never been tested in practice* by comparing it with other approaches. However, articles [183, 186] were published in 2005 and in 2004, respectively, and in 2009 the Dutch authors proposed a methodology for weighing the criteria and successfully tested the validity of using this technique for known carcinogens (G. Swaen and L. van Amelsvoort [67]). We plan to consider data of this kind on the gradation and weighting of criteria, as well as other approaches to establishing causation [187], in Report 4. We should only note that Rothman and Greenland [42, 43, 99], as well as Goodman and Phillips [188, 189], attempted to propose such approaches (in our subjective opinion, only in theory).

Summing up this section, we will formulate only the following indisputable statement: the criterion “Biological Plausibility” is not absolute. It is neither necessary, despite all its importance, nor sufficient [11, 13, 27, 29, 38, 49, 58, 64], despite its role in the set of rules of causality for environmental protection and toxicology agencies [16]. And yet, without it, any proof appears both incomplete and flawed. Something like a relationship between ice cream consumption and drowning deaths [1] or between the linear size of the penis and the reciprocal of the country’s GDP [80] will always come to mind.

## CONCLUSIONS

In this section, the main array of references is not provided; they can be found above.

In the absence of direct experiments, biological plausibility has been an important component of causality hypotheses in the biomedical disciplines since at least the 19th century. During the formation of the guidelines/criteria for assessing the plausibility of a causal relationship between exposure and effect (1950–1960s), the point “Biological Plausibility” was one of the first to be proposed (Lilienfeld in 1957). Later, this statement was sometimes replaced with the similar criterion “Coherence” (agreement with the known facts of natural history and biology of the disease, 1964); however, in 1965 Hill introduced both points into his list.

Since then, the guideline “Biological Plausibility” has gained increasing importance, especially in disciplines (ecology, toxicology, the study of carcinogens) in which there are difficulties not only in setting up adequate experiments, but even in the very observation of the effect. In such cases, to assess the probable effects of exposure for subsequent decision-making in the healthcare sector, one has also to rely on molecular epidemiology (“surrogate EndPoints,” biomarkers),



as well as animal and in vitro experiments. This situation is observed, among other things, in US courts when making decisions on compensation for various impacts ("Daubert's rule").

The criterion "Biological Plausibility" remains important in classical epidemiology; adherence to it is highly desirable and sometimes necessary even when the correlations are statistically established. Epidemiological approaches alone are not able to prove the true causality of the relationship (possibly the influence of chance, confounders, biases, and reverse causality). Statistical approaches are aimed only at proving the reality and persistence of associations. Without knowledge of the biological mechanism and plausible model, such correlations (especially weak ones) cannot be regarded as confirmation of a proposed hypothesis about the true causality of the effect of exposure.

The essence of this criterion is reduced to the integration of data from various disciplines of medical and biological profiles. This approach is so important in establishing causality that authors specifically emphasize its necessity in a number of sources, in specialized publications, and in textbooks on epidemiology. This review provides relevant citations, including for radiobiology and radiation epidemiology. It turns out that data from almost every area of the biomedical and molecular-cellular disciplines can make an important contribution to the search for evidence for epidemiological research. This imparts practical significance to almost any experimental research, no matter how fundamental and theoretical it may seem.

Several leading authors have detailed the levels of attaining biological plausibility. In the work by Weed and Hursting from 1998 [29], three levels are proposed, from the first, when there is an assumption about the mechanism, but there is no biological evidence, through the second, when evidence is provided by molecular biology and molecular epidemiology, to the third, highest level, when there are data on the influence of a particular factor of interest on the mechanism. Similarly, Susser in 1977 and 1986 [25, 36, 37] formulated four levels for the "Coherence" criterion with a similar meaning. Attaining its significance includes the following elements of consistency with (1) theoretical plausibility, (2) biological knowledge (i.e., "Biological Plausibility"), (3) facts, and (4) statistical patterns, including the dose-effect dependence.

Ultimately, it is all a matter of judgment: is there enough evidence to rule out alternative explanations ("Weight of Evidence"). In addition, the decision on causality is largely based on the intuition of researchers [33, 163], which was mentioned by Hill in 1965

[11]. The last statement fully applies to attaining an acceptable level of biological plausibility. However, the aforementioned Phillips and Goodman in 2004 [190] noted that the correct approach to intuitive assessment is now almost completely lost. For example, studies performed in the 1970s and 1980s demonstrated that both ordinary people and experts have poor quantitative intuition and, therefore, intuition cannot eliminate the need for modern methods of quantitative estimation of uncertainty [190].

One of the well-known methodologies for integrating data from different disciplines is Bayesian analysis (Bayesian meta-analysis). It relies on the Weight of Evidence (proofs) rather than on a certain specific study, following the same principles as Bayesian decision-making. This method implies the integration of all available biomedical data into one updated and standardized estimate that determines the overall strength of causation.

Regardless of the importance and necessity of the "Biological Plausibility" criterion, it, similarly to almost all of Hill's criteria (eight out of nine, except for Temporality), is neither necessary nor sufficient for evidence. There are a number of examples (including radiation exposure) given in this publication that show, first, that "plausibility" depends on the current biological knowledge and, second, that there are real but seemingly implausible associations, and vice versa.

From this follows the criticism of both the "Biological Plausibility" criterion as such and the entire inductive approach based on causal criteria by some authors. They point to the insufficient scientificity and accuracy of the approach (Feinstein), subjectivity, dogmatism, and non-absoluteness of criteria (Rothman and Greenland), to the incorrectness of methods that integrate data from different disciplines that are not the same in approach, methodology, or evidence (Charlton), and to the fact that the criteria of causality have never been tested in practice (Goodman and Phillips). The last statement was made in 2005; however, since then the situation has changed: in 2009, Hill's criteria were modified by determining the weight of each, and the methodology was tested on standard carcinogenic agents (Swaen and van Amelsvoort [67]).

Despite the expressed criticism and doubts and despite the fact that the criterion "Biological Plausibility" was said to be not absolute, any epidemiological evidence without it looks both incomplete and flawed. As noted in [191], the fact that epidemiology is a biological science related to human diseases and based on clinical disciplines cannot be neglected.

This publication, together with the previous special reviews on the "Strength of Association" [4, 5] and

“Temporality” [3], constitutes a detailed preamble to the planned Reports 3 and 4 of our cycle on the essence, limitations, breadth of application, and radiation aspect of the guiding principles/criteria for causality.

#### POSSIBILITY OF SUBJECTIVE BIAS

This study, carried out within the framework of the broader budgetary theme of R&D of the Federal Medical Biological Agency of Russia, was not supported by any other sources of funding. There were no time frames, official requirements, restrictions, or other external objective or subjective confounding factors in the performance of this study.

#### COMPLIANCE WITH ETHICAL STANDARDS

The authors declare that they have no conflict of interest. This article does not contain any studies involving animals or human participants performed by any of the authors.

#### REFERENCES

- Koterov, A.N., Causal criteria in medical and biological disciplines: history, essence and radiation aspect, Report 1: Problem statement, conception of causes and causation, *Radiats. Biol., Radioekol.*, 2019, vol. 59, no. 1, pp. 1–32. <https://doi.org/10.1134/S0869803119010065>
- Koterov, A.N., Causal criteria in medical and biological disciplines: history, essence and radiation aspect, Report 2: Henle–Koch postulates and criteria for causality of non-communicable pathologies before Hill, *Radiats. Biol., Radioekol.*, 2019, vol. 59, no. 4, pp. 341–375. <https://doi.org/10.1134/S0869803119040052>
- Koterov, A.N., Ushenkova, L.N., and Biryukov, A.P., Hill’s criteria “Temporality,” Reverse causation and its radiation aspect, *Radiats. Biol., Radioekol.*, 2020, vol. 60, no. 2, pp. 115–152.
- Koterov, A.N., Ushenkova, L.N., Zubenkova, E.S., et al., Strength of association, Report 1: Graduation of relative risk, *Med. Radiol. Radiats. Bezop.*, 2019, vol. 64, no. 4, pp. 5–17. [https://doi.org/10.12737/article\\_5d1adb25725023.14868717](https://doi.org/10.12737/article_5d1adb25725023.14868717)
- Koterov, A.N., Ushenkova, L.N., Zubenkova, E.S., et al., Strength of association, Report 2: Graduation of correlation size, *Med. Radiol. Radiats. Bezop.*, 2019, vol. 64, no. 6, pp. 12–24.
- Susser, M., *Causal Thinking in the Health Sciences: Concepts and Strategies of Epidemiology*, New York: Oxford Univ. Press, 1973.
- Ecological Causal Assessment*, Norton, S.B., Cormier, S.M., and Suter, G.W., II, Eds., Cincinnati, OH: CRC Press, 2015.
- Health Risks from Exposure to Low Levels of Ionizing Radiation: BEIR VII Phase 2*, Washington, DC: National Academies Press, 2006. <http://www.nap.edu/catalog/11340.html>. Cited July 4, 2019.
- Doll, R., Weak associations in epidemiology: importance, detection, and interpretation, *J. Epidemiol.*, 1996, vol. 6, no. 4, pp. S11–S20. [https://doi.org/10.2188/jea.6.4sup\\_11](https://doi.org/10.2188/jea.6.4sup_11)
- Causality in the Sciences*, Illari, P.M., Russo, F., and Williamson, J., Eds., New York: Oxford Univ. Press, 2011. <https://doi.org/10.1093/acprof:oso/9780199574131.001.0001>
- Hill, B.A., The environment and disease: association or causation? *Proc. R. Soc. Med.*, 1965, vol. 58, no. 5, pp. 295–300. <https://doi.org/10.1177/0141076814562718>
- Evans, A.S., Causation and disease: the Henle-Koch postulates revisited, *Yale J. Biol. Med.*, 1976, vol. 49, no. 2, pp. 175–195.
- Glynn, J.R., A question of attribution, *Lancet*, 1993, vol. 342, no. 8870, pp. 530–532. [https://doi.org/10.1016/0140-6736\(93\)91651-2](https://doi.org/10.1016/0140-6736(93)91651-2)
- A Dictionary of Epidemiology*, Last, J.M., Ed., Oxford: Oxford Univ. Press, 2001.
- Fox, G.A., Practical causal inference for ecoepidemiologists, *J. Toxicol. Environ. Health*, 1991, vol. 33, no. 4, pp. 359–273. <https://doi.org/10.1080/15287399109531535>
- Meek, M.E., Palermo, C.M., Bachman, A.N., et al., Mode of action human relevance (species concordance) framework: evolution of the Bradford Hill considerations and comparative analysis of weight of evidence, *J. Appl. Toxicol.*, 2014, vol. 34, no. 6, pp. 595–606. <https://doi.org/10.1002/jat.2984>
- Weed, D.L. and Gorelic, L.S., The practice of causal inference in cancer epidemiology, *Cancer Epidemiol. Biomarkers Prev.*, 1996, vol. 5, no. 4, pp. 303–311.
- Wynder, E.L., Environmental causes of cancer in man, *Med. Clin. North Am.*, 1956, vol. 40, no. 3, pp. 624–643. [https://doi.org/10.1016/S0025-7125\(16\)34556-4](https://doi.org/10.1016/S0025-7125(16)34556-4)
- Wynder, E.L. and Day, E., Some thoughts on the causation of chronic disease, *J. Am. Med. Assoc.*, 1961, vol. 175, no. 11, pp. 997–999. <https://doi.org/10.1001/jama.1961.63040110004014>
- Lilienfeld, A.M., Epidemiological methods and inferences in studies of non-infectious diseases, *Publ. Health Rep.*, 1957, vol. 72, no. 1, pp. 51–60. <https://doi.org/10.2307/4589685>
- Lilienfeld, A.M., On the methodology of investigations of etiologic factors in chronic diseases: some comments, *J. Chronic Dis.*, 1959, vol. 10, no. 1, pp. 41–46. [https://doi.org/10.1016/0021-9681\(59\)90016-5](https://doi.org/10.1016/0021-9681(59)90016-5)
- Sartwell, P.E., “On the methodology of investigations of etiologic factors in chronic diseases”—Further comments, *J. Chronic Dis.*, 1960, vol. 11, no. 1, pp. 61–63. [https://doi.org/10.1016/0021-9681\(60\)90140-5](https://doi.org/10.1016/0021-9681(60)90140-5)
- Smoking and Health: Report of the Advisory Committee to the Surgeon General of the Public Health Service Publication No. 1103*, Washington, DC: US Dep. Health, Educ. Welfare, 1964. <https://profiles.nlm.nih.gov/ps/access/nnbbmq.pdf>. Cited July 4, 2019.
- Vlasov, V.V., *Epidemiologiya (Epidemiology)*, Moscow: GEOTAR-Media, 2006, 2nd ed.

25. Susser, M., Judgement and causal inference: criteria in epidemiologic studies, *Am. J. Epidemiol.*, 1995, vol. 141, no. 8, pp. 701–715.
26. Thygesen, L.C., Andersen, G.S., and Andersen, H., A philosophical analysis of the Hill criteria, *J. Epidemiol. Comm. Health*, 2005, vol. 59, no. 6, pp. 512–516. <https://doi.org/10.1136/jech.2004.027524>
27. Strom, B.L., Study designs available for pharmacoepidemiology studies, in *Textbook of Pharmacoepidemiology*, Strom B.L., Ed., Chichester: Wiley, 2000, ch. 2, pp. 17–30.
28. Friis, R.H. and Sellers, T.A., *Epidemiology for Public Health Practice*, 5th ed., Burlington: Jones & Bartlett Learning, 2014.
29. Weed, D.L. and Hursting, S.D., Biologic plausibility in causal inference: current method and practice, *Am. J. Epidemiol.*, 1998, vol. 147, no. 5, pp. 415–425. <https://doi.org/10.1093/oxfordjournals.aje.a009466>
30. Goodman, S.N. and Samet, J.M., Cause and cancer epidemiology, in *Schottenfeld and Fraumeni Cancer Epidemiology and Prevention*, Thun, M.J., , Eds., 4th ed., New York: Oxford Univ. Press, 2018, pp. 97–104.
31. Egilman, D., Kim, J., and Biklen, M., Proving causation: the use and abuse of medical and scientific evidence inside the courtroom—an epidemiologist's critique of the judicial interpretation of the Daubert ruling, *Food Drug Law J.*, 2003, vol. 58, no. 2, pp. 223–250.
32. Hollingsworth, J.G. and Lasker, E.G., The case against differential diagnosis: Daubert, medical causation testimony, and the scientific method, *J. Health Law*, 2004, vol. 37, no. 1, pp. 85–111.
33. Bhopal, R.S., *Concepts of Epidemiology: Integrating the Ideas, Theories, Principles, and Methods of Epidemiology*, Oxford: Oxford Univ. Press, 2016, 3rd ed.
34. Fedak, K.M., Bernal, A., Capshaw, Z.A., and Gross, S., Applying the Bradford Hill criteria in the 21st century: how data integration has changed causal inference in molecular epidemiology, *Emerg. Themes Epidemiol.*, 2015, vol. 12, art. 14. <https://doi.org/10.1186/s12982-015-0037-4>
35. Gordis L., *Epidemiology*, Philadelphia: Saunders, 2014, 5th ed.
36. Susser, M., Rules of inference in epidemiology, *Regul. Toxicol. Pharmacol.*, 1986, vol. 6, no. 2, pp. 116–128. [https://doi.org/10.1016/0273-2300\(86\)90029-2](https://doi.org/10.1016/0273-2300(86)90029-2)
37. Susser, M., The logic of Sir Karl Popper and the practice of epidemiology, *Am. J. Epidemiol.*, 1986, vol. 124, no. 5, pp. 711–718. <https://doi.org/10.1093/oxfordjournals.aje.a114446>
38. *Guidelines for Carcinogen Risk Assessment, EPA/630/P-03/001B*, Washington, DC: US Environ. Prot. Agency, 2005.
39. Weed, D.L., Epidemiologic evidence and causal inference, *Hematol. Oncol. Clin. North Am.*, 2000, vol. 14, no. 4, pp. 797–807. [https://doi.org/10.1016/S0889-8588\(05\)70312-9](https://doi.org/10.1016/S0889-8588(05)70312-9)
40. Becker, R.A., Dellarco, V., Seed, J., et al., Quantitative weight of evidence to assess confidence in potential modes of action, *Regul. Toxicol. Pharmacol.*, 2017, vol. 86, pp. 205–220. <https://doi.org/10.1016/j.yrtph.2017.02.017>
41. Rothman, K., *Modern Epidemiology*, Boston: Little Brown, 1986, 1st ed.
42. Rothman, K.J., *Modern Epidemiology*, Philadelphia: Lippincott Williams and Wilkins, 1998, 2nd ed.
43. Rothman, K.J., Greenland, S., Poole, C., and Lash, T.L., Causation and causal inference, in *Modern Epidemiology*, Rothman, K.J., Greenland, S., and Lash, T.L., Eds., 3rd ed., Philadelphia, PA: Kluwer, 2008, pp. 5–31.
44. Rothman, K.J., Inferring causal connections—habit, faith or logic? in *Causal Inference*, Rothman, K.J., Ed., Chestnut Hill, MA: Epidemiology Resources, 1988, pp. 3–12.
45. Rothman, K. and Greenland, S., Hill's criteria for causality, in *Encyclopedia of Biostatistics*, Chichester: Wiley, 2005. <https://www.rtihs.org/sites/default/files/26902%20Rothman%201998%20The%20encyclopedia%20of%20biostatistics.pdf>. Cited July 4, 2019.
46. Merrill, R.M., Frankenfeld C.L., Freeborne N., and Mink M., *Behavioral Epidemiology. Principles and Applications*, Burlington: Jones & Bartlett Learning, 2016.
47. Lagiou P, Adami H.O., and Trichopoulos D., Causality in cancer epidemiology, *Eur. J. Epidemiol.*, 2005, vol. 20, no. 7, pp. 565–574. <https://doi.org/10.1007/sl0654-005-7968-y>
48. Aschengrau, A. and Seage, G.R. III, *Epidemiology in Public Health*, Burlington: Jones & Bartlett Learning, 2014, 3rd ed.
49. *Smoking and Reproductive Life. The Impact of Smoking on Sexual, Reproductive and Child Health*, Carter, D., Nathanson, N., Seddon, C., et al., Eds., London: Br. Med. Assoc., 2004. [https://www.rauchfrei-info.de/fileadmin/main/data/Dokumente/Smoking\\_ReproductiveLife.pdf](https://www.rauchfrei-info.de/fileadmin/main/data/Dokumente/Smoking_ReproductiveLife.pdf). Cited July 4, 2019.
50. *Guidelines for Carcinogen Risk Assessment, EPA/630/R-95/002F*, Washington, DC: US Environ. Prot. Agency, 1998.
51. Weed, D.L., Precaution, prevention, and public health ethics, *J. Med. Philos.*, 2004, vol. 29, no. 3, pp. 313–332. <https://doi.org/10.1080/03605310490500527>
52. Porta, M., Biologic plausibility in causal inference: current method and practice, *Am. J. Epidemiol.*, 1999, vol. 150, no. 2, pp. 217–219. <https://doi.org/10.1093/oxfordjournals.aje.a009984>
53. Carbone, M., Klein, G., Gruber, J., and Wong, M., Modern criteria to establish human cancer etiology, *Cancer Res.*, 2004, vol. 64, no. 15, pp. 5518–5524. <https://doi.org/10.1158/0008-5472.CAN-04-0255>
54. Russo, F. and Williamson, J., Interpreting causality in the health sciences, *Int. Stud. Philos. Sci.*, 2007, vol. 21, no. 2, pp. 157–170. <https://doi.org/10.1080/02698590701498084>
55. Cormier, S.M., Suter, G.W. II, and Norton, S.B., Causal characteristics for ecoepidemiology, *Hum. Ecol. Risk Assess.*, 2010, vol. 16, no. 1, pp. 53–73. <https://doi.org/10.1080/10807030903459320>
56. Hartung, T., Luechtefeld, T., Maertens, A., and Kleensang, A., Integrated testing strategies for safety assessments, *ALTEX*, 2013, vol. 30, no. 1, pp. 3–18. <https://doi.org/10.14573/altex.2013.1.003>

57. *Report of the Advisory Group to Plan, Vol. 100: A Review of Human Carcinogenesis: IARC Internal Report No. 07/001*, Lyon: Int. Agency Res. Cancer, 2007.
58. Fletcher, R.H., Fletcher, S.W., and Wagner, E.H., *Clinical Epidemiology: The Essentials*, Philadelphia, PA: Lippincott Williams & Wilkins, 1996.
59. Weed, D.L., On the use of causal criteria, *Int. J. Epidemiol.*, 1997, vol. 26, no. 6, pp. 1137–1141. <https://doi.org/10.1093/ije/26.6.1137>
60. Kaufman, J.S. and Poole, C., Looking back on “causal thinking in the health sciences”, *Annu. Rev. Publ. Health*, 2000, vol. 21, pp. 101–119. <https://doi.org/10.1146/annurev.publhealth.21.1.101>
61. Ward, A.C., The role of causal criteria in causal inferences: Bradford Hill’s “aspects of association,” *Epidemiol. Perspect. Innovations*, 2009, vol. 6, no. 2. <https://doi.org/10.1186/1742-5573-6-2>
62. Bruce, N., Pope, D., and Stanistreet, D., *Quantitative Methods for Health Research. A Practical Interactive Guide to Epidemiology and Statistics*, Chichester: Wiley, 2019, 2nd ed.
63. Merrill, R.M., *Introduction to Epidemiology*, Burlington: Jones & Bartlett Learning, 2017, 7th ed.
64. Webb, P. and Bain, C., *Essential Epidemiology: An Introduction for Students and Health Professionals*, Cambridge: Cambridge Univ. Press, 2011, 2nd ed.
65. Greenland, S., Relation of probability of causation to relative risk and doubling dose: a methodologic error that has become a social problem, *Am. J. Publ. Health*, 1999, vol. 89, no. 8, pp. 1166–1169. <https://doi.org/10.2105/AJPH.89.8.1166>
66. Moraes, R., Gerhard, P., Andersson, L., et al., Establishing causality between exposure to metals and effects on fish, *Hum. Ecol. Risk Assess.*, 2003, vol. 9, no. 1, pp. 149–169. <https://doi.org/10.1080/713609857>
67. Swaen, G. and van Amelsvoort, L., A weight of evidence approach to causal inference, *J. Clin. Epidemiol.*, 2009, vol. 62, no. 3, pp. 270–277. <https://doi.org/10.1016/j.jclinepi.2008.06.013>
68. Kundi, M., Causality and the interpretation of epidemiologic evidence, *Environ. Health Perspect.*, 2006, vol. 114, no. 7, pp. 969–974. <https://doi.org/10.1289/ehp.8297>
69. Elwood, J.M., *Causal Relationships in Medicine—a Practical System for Critical Appraisal*, Oxford: Oxford Univ. Press, 1988.
70. Renton, A., Epidemiology and causation: a realist view, *J. Epidemiol. Community Health*, 1994, vol. 48, no. 1, pp. 79–85. <https://doi.org/10.1136/jech.48.1.79>
71. Susser, M., What is a cause and how do we know one? A grammar for pragmatic epidemiology, *Am. J. Epidemiol.*, 1991, vol. 133, no. 7, pp. 635–648. <https://doi.org/10.1093/oxfordjournals.aje.a115939>
72. Angell, M., The interpretation of epidemiologic studies, *N. Engl. J. Med.*, 1990, vol. 323, no. 12, pp. 823–825. <https://doi.org/10.1056/NEJM199009203231209>
73. Coughlin, S.S., *Causal Inference and Scientific Paradigms in Epidemiology*, Sharjah: Bentham Science, 2010. <https://doi.org/10.2174/97816080518161100101>. <https://ebooks.bentham-science.com/book/9781608051816/>. Cited July 4, 2019.
74. Rothman, K.J., *Epidemiology. An Introduction*, Oxford: Oxford Univ. Press, 2012, 2nd ed.
75. Lower, G.M. and Kanarek, M.S., Conceptual/operational criteria of causality: relevance to systematic epidemiologic theory, *Med. Hypotheses*, 1983, vol. 11, no. 2, pp. 217–244. [https://doi.org/10.1016/0306-9877\(83\)90064-6](https://doi.org/10.1016/0306-9877(83)90064-6)
76. Martin, C.J., Sutton, D.G., West, C.M., and Wright, E.G., The radiobiology/radiation protection interface in healthcare, *J. Radiol. Prot.*, 2009, vol. 29, no. 2, pp. A1–A20. <https://doi.org/10.1088/0952-4746/29/2A/S01>
77. Trott, K.R. and Rosemann, M., Molecular mechanisms of radiation carcinogenesis and the linear, non-threshold dose response model of radiation risk estimation, *Radiat. Environ. Biophys.*, 2000, vol. 39, no. 2, pp. 79–87. <https://doi.org/10.1007/s004110000047>
78. Koterov, A.N., Zharkova, G.P., and Biryukov, A.P., Tandem of radiation epidemiology and radiobiology for practice and radiation protection, *Med. Radiol. Radiats. Bezop.*, 2010, vol. 55, no. 4, pp. 55–84.
79. Preston, R.J., Integrating basic radiobiological science and epidemiological studies: why and how, *Health Phys.*, 2015, vol. 108, no. 2, pp. 125–130.
80. Westling, T., Male organ and economic growth: does size matter?, 2011. <https://helda.helsinki.fi/bitstream/handle/10138/27239/maleorga.pdf>. Cited July 4, 2019.
81. Eddy, D.M., The confidence profile method: a Bayesian method for assessing health technologies, *Oper. Res.*, 1989, vol. 37, no. 2, pp. 210–228. <https://doi.org/10.1287/opre.37.2.210>
82. Hoaglin, D.C., Hawkins, N., Jansen, J.P., et al., Conducting indirect treatment comparisons and network meta-analysis for health care decision making: report of the ISPOR task force on indirect treatment comparisons good research practices: Part 2, *Value Health*, 2011, vol. 14, pp. 429–437. <https://doi.org/10.1016/j.jval.2011.01.011>
83. Drivotinov, B.V., Apanel’, E.N., and Mastyskin, A.S., The basic terms accepted in evidence based medicine, *Med. Zh.* (Minsk), 2005, no. 2, pp. 103–107. <http://rep.bsmu.by/handle/BSMU/5255>. Cited July 4, 2019.
84. Lewis, M.G. and Nair, N.S., Review of applications of Bayesian meta-analysis in systematic reviews, *Global J. Med. Publ. Health*, 2015, vol. 4, no. 1. <http://www.gjmedph.com/uploads/R1-Vo4No1.pdf>. Cited July 4, 2019.
85. Blettner, M., Sauerbrei, W., Schlehofer, B., et al., Traditional reviews, meta-analyses and pooled analyses in epidemiology, *Int. J. Epidemiol.*, 1999, vol. 28, no. 1, pp. 1–9. <https://doi.org/10.1093/ije/28.1.1>
86. Wald, N.J. and Morris, J.K., Teleoanalysis: combining data from different types of study, *Br. Med. J.*, 2003,

- vol. 327, no. 7415, pp. 616–618.  
<https://doi.org/10.1136/bmj.327.7415.616>
87. Howick, J., Glasziou, P., and Aronson, J.K., The evolution of evidence hierarchies: what can Bradford Hill's 'guidelines for causation' contribute? *J. R. Soc. Med.*, 2009, vol. 102, no. 5, pp. 186–194.  
<https://doi.org/10.1258/jrsm.2009.090020>
  88. Aronson, J.K., Unity from diversity: the evidential use of anecdotal reports of adverse drug reactions and interactions, *J. Eval. Clin. Pract.*, 2005, vol. 11, no. 2, pp. 195–208.  
<https://doi.org/10.1111/j.1365-2753.2005.00527.x>
  89. Loke, Y.K., Adverse drug reactions, *Br. J. Clin. Pharmacol.*, 2012, vol. 73, no. 6, pp. 908–911.  
<https://doi.org/10.1111/j.1365-2125.2012.04235.x>
  90. Howick, J., Glasziou, P., and Aronson, J.K., Evidence-based mechanistic reasoning, *J. R. Soc. Med.*, 2010, vol. 103, no. 11, pp. 433–441.  
<https://doi.org/10.1258/jrsm.2010.100146>
  91. Guzelian, P.S., Victoroff, M.S., Halmes, N.C., et al., Evidence-based toxicology: a comprehensive framework for causation, *Hum. Exp. Toxicol.*, 2005, vol. 24, no. 4, pp. 161–201.  
<https://doi.org/10.1191/0960327105ht517oa>
  92. *The Health Consequences of Smoking: A Report of the Surgeon General*, Rockville, MD: Off. Surg. Gen., US Publ. Health Serv., 2004.
  93. *Handbook of Epidemiology*, Ahrens, W. and Pigeot, I., Eds., New York: Springer-Verlag, 2014, 2nd ed.
  94. Farr, W., The mortality of cholera in England, 1848–49, in *Vital Statistics: A Memorial Volume of Selections from the Reports and Writings of William Farr*, Humphreys, N.A., Ed., London: Sanitary Inst., 1885, pp. 333–351.
  95. Snow, J., *On the Mode of Communication of Cholera*, London: John Churchill, 1855.
  96. Szklo, M. and Nieto, F.J., *Epidemiology: Beyond the Basics*, Burlington: Jones & Bartlett Learning, 2019, 4th ed.
  97. Bonita, R., Beaglehole, R., and Kjellstrom, T., *Basic Epidemiology*, Geneva: World Health Org., 2006, 2nd ed.
  98. Cheever, D.W., The value and fallacy of statistics in the observation of disease, *Boston Med. Surg. J.*, 1861, vol. 63, no. 24, pp. 449–456.  
<https://doi.org/10.1056/nejm186101240632609>
  99. Rothman, K.J. and Greenland, S., Causation and causal inference in epidemiology, *Am. J. Publ. Health*, 2005, vol. 95, suppl. 1, pp. S144–S150.  
<https://doi.org/10.2105/AJPH.2004.059204>
  100. Phillips, A.N. and Davey Smith, G., Confounding in epidemiological studies, *Br. Med. J.*, 1993, vol. 306, no. 870, p. 142.  
<https://doi.org/10.1136/bmj.306.6870.142-b>
  101. Glaser, H., *Dramatische Medizin: Selbstversuche von Ärzten*, Zurich: Orell Furssli, 1959.
  102. *Closing the Circle on the Splitting of the Atom: The Environmental Legacy of Nuclear Weapons Production in the United States and What the Department of Energy Is Doing About It*, Washington, DC: Environ. Law Inst., 1995.  
[https://www.energy.gov/sites/prod/files/2014/03/f8/Closing\\_the\\_Circle\\_Report.pdf](https://www.energy.gov/sites/prod/files/2014/03/f8/Closing_the_Circle_Report.pdf). Cited July 4, 2019.
  103. Stewart, F.A., Akleyev, A.V., Hauer-Jensen, M., Hendry, J.H., Kleiman, N.J., MacVittie, T.J., Aleman, B.M., Edgar, A.B., Mabuchi, K., Muirhead, C.R., Shore, R.E., and Wallace, W.H., ICRP publication 118: ICRP statement on tissue reactions and early and late effects of radiation in normal tissues and organs—threshold doses for tissue reactions in a radiation protection context, *Ann. ICRP*, 2012, vol. 41, nos. 1–2, pp. 1–322.
  104. Clement, C., *The International Framework for Radiation Safety. History, Science, Philosophy and Practice*, Dublin: Radiol. Prot. Inst. Ireland, 2009.  
[https://www.epa.ie/pubs/reports/radiation/RPII\\_I-CRP\\_Int\\_Frame\\_Rad\\_09.pdf](https://www.epa.ie/pubs/reports/radiation/RPII_I-CRP_Int_Frame_Rad_09.pdf). Cited July 4, 2019.
  105. Tubiana, M., Radiation risks in perspective: radio-induced cancer among cancer risks, *Proc. First Int. Symp. "The Effects of Low and Very Low Dose of Ionizing Radiation on Human Health"*, Amsterdam: Elsevier, 2000, pp. 3–24.
  106. Macklis, R.M., Radithor and the era of mild radium therapy, *J.A.M.A.*, 1990, vol. 264, no. 5, pp. 614–618.  
<https://doi.org/10.1001/jama.1990.03450050072031>
  107. Jorgensen, T.J., *Strange Glow. The Story of Radiation*, Princeton: Princeton Univ. Press, 2016.
  108. Brodsky, A., Kathren, R.L., and Willis, C.A., History of the medical uses of radiation: regulatory and voluntary standards of protection, *Health Phys.*, 1995, vol. 69, no. 5, pp. 783–823.  
<https://doi.org/10.1097/00004032-199511000-00012>
  109. Mottram, J.C. and Russ, S., Lymphopenia following exposures of rats to "soft" X-rays and the beta-rays of radium, *J. Exp. Med.*, 1921, vol. 34, no. 3, pp. 271–273.  
<https://doi.org/10.1084/jem.34.3.271>
  110. Sugiura, K. and Failla, G., Some effects of radium radiations on white mice, *J. Gen. Physiol.*, 1922, vol. 4, no. 4, pp. 423–436.  
<https://doi.org/10.1085/jgp.4.4.423>
  111. Lazarus-Barlow, W.S., On the disappearance of insoluble radium salts from the bodies of mice after subcutaneous injection, *Arch. Middlesex Hosp.*, 1913, vol. 30, pp. 92–94.
  112. Lazarus-Barlow, W.C., The causation of cancer, *Br. Med. J.*, 1924, vol. 1, no. 3301, pp. 647–648.
  113. Orr, C.P. and Eben, M., Byers: the effect of gamma rays on amateur golf, modern medicine and the FDA, *Heritage. Allegheny Cem.*, 2004, vol. 13, no. 1, pp. 6–7.  
[http://www.allegheny cemetery.com/images/newsletter/newsletter\\_XIII\\_1.pdf](http://www.allegheny cemetery.com/images/newsletter/newsletter_XIII_1.pdf). Cited July 1, 2019.
  114. Lock, stock, and history, Vita Radium Suppositories, Pinterest.  
<https://www.orau.org/ptp/collection/quackcures/radsup.htm>. Cited July 4, 2019.
  115. Vita Radium Suppositories (ca.1930), Produced by the Home Products Company of Denver, Colorado.  
<https://www.orau.org/ptp/collection/quackcures/radsup.htm>. Cited July 4, 2019.
  116. Belozerskii, G.N., *Radiatsionnaya ekologiya: Uchebnik dlya bakalavriata i magistratury* (Radiation Ecology: Manual for Bachelor's and Master's Degree Students), Moscow: Yurait, 2019, 2nd ed.
  117. Frame, P. and Kolb, W., *Living With Radiation: The First Hundred Years*, Baltimore, MD: Syntec, 2000,

- 2nd ed. <https://www.orau.org/ptp/collection/atomic-toys/atomic-toys.htm>. Cited July 4, 2019.
118. *Report to the General Assembly, with Scientific Annexes*, New York: UN Sci. Com. Effects At. Radiat., 2001, annex C, pp. 5–160.
  119. Koterov, A.N., The main question of radiation genetics: Do people have radiation mutants? in *Aktual'naya radiobiologiya: Kurs lektsii* (Actual Radiobiology: A Course of Lectures), Smirnov, V.P., et al., Ed., Moscow: Mosk. Energ. Inst., 2015, no. 4, pp. 133–158.
  120. Koterov, A.N. and Biryukov, A.P., The offspring of liquidators of Chernobyl Atomic Power Station accident. 1. The estimation of the basic opportunity to register of radiation effect, *Med. Radiol. Radiats. Bezop.*, 2012, vol. 57, no. 1, pp. 58–79.
  121. Koterov, A.N. and Biryukov, A.P., The offspring of liquidators of Chernobyl Atomic Power Station accident. 2. The frequency of anomalies and pathologies and its connection to non-radiation factors, *Med. Radiol. Radiats. Bezop.*, 2012, vol. 57, no. 2, pp. 51–77.
  122. ICRP Publication 103: P103: The 2007 recommendations of the International Commission on Radiological Protection. *Ann. ICRP*, 2007, vol. 37, nos. 2–4.
  123. Zakharova, M.L., Bezlepkin, V.G., Kirillova, E.N., et al., Radiobiology human tissue repository genetic material and the certain results of its study, *Med. Radiol. Radiats. Bezop.*, 2010, vol. 55, no. 5, pp. 5–13.
  124. Bezlepkin, V.G., Kirillova, E.N., Zakharova, M.L., et al., Delayed and transgenerational molecular and genetic effects of prolonged influence of ionizing radiation in nuclear plant workers, *Radiats. Biol., Radioekol.*, 2011, vol. 51, no. 1, pp. 20–32.
  125. Kuzmina, N.S., Myazin, A.E., Lapteva, N.Sh., and Rubanovich, A.V., The study of hypermethylation in irradiated parents and their children blood leukocytes, *Cent. Eur. J. Biol.*, 2014, vol. 9, no. 10, pp. 941–950. <https://doi.org/10.2478/s11535-014-0333-y>
  126. Kuzmina, N.S., Lapteva, N.Sh., and Rubanovich, A.V., Hypermethylation of gene promoters in peripheral blood leukocytes in humans long term after radiation exposure, *Environ. Res.*, 2016, vol. 146, pp. 10–17. <https://doi.org/10.1016/j.envres.2015.12.008>
  127. Kuzmina, N.S., Lapteva, N.Sh., Rusinova, G.G., et al., Gene hypermethylation in blood leukocytes in humans long term after radiation exposure. Validation set, *Environ. Pollut.*, 2018, vol. 234, pp. 935–942. <https://doi.org/10.1016/j.envpol.2017.12.039>
  128. Diggle, G.E., Thalidomide: 40 years on, *Int. J. Clin. Pract.*, 2001, vol. 55, no. 9, pp. 627–631.
  129. Jacobson, R.M. and Feinstein, A.R., Oxygen as a cause of blindness in premature infants: “autopsy” of a decade of errors in clinical epidemiologic research, *J. Clin. Epidemiol.*, 1992, vol. 45, no. 11, pp. 1265–1287.
  130. Lacchetti, C., Ioannidis, J., and Guyatt, G., Surprising results of randomized trials, in *Users' Guides to the Medical Literature: A Manual for Evidence-Based Clinical Practice*, Guyatt, G., Rennie, D., Meade, M.O., and Cook, D.J., Eds., 2nd ed., New York: McGraw Hill, 2008, pp. 113–151.
  131. Howick, J., *The Philosophy of Evidence-Based Medicine*, Chichester: Wiley-Blackwell, 2011.
  132. Nguyen, T.K., Nguyen, E.K., Warner, A., et al., Failed randomized clinical trials in radiation oncology: what can we learn? *Int. J. Radiat. Oncol. Biol. Phys.*, 2018, vol. 101, no. 5, pp. 1018–1024. <https://doi.org/10.1016/j.ijrobp.2018.04.030>
  133. Krauss, A., Why all randomized controlled trials produce biased results, *Ann Med.*, 2018, vol. 50, no. 4, pp. 312–322. <https://doi.org/10.1080/07853890.2018.1453233>
  134. Cartwright, N., Are RCTs the gold standard? *Biosocieties*, 2007, vol. 2, no. 1, pp. 11–20. <https://doi.org/10.1017/S1745855207005029>
  135. Cartwright, N. and Munro, E., The limitations of randomized controlled trials in predicting effectiveness, *J. Eval. Clin. Pract.*, 2010, vol. 16, no. 2, pp. 260–266. <https://doi.org/10.1111/j.1365-2753.2010.01382.x>
  136. Grossman, J. and Mackenzie, F., The randomized controlled trial: gold standard, or merely standard? *Perspect. Biol. Med.*, 2005, vol. 48, no. 4, pp. 516–534. <https://doi.org/10.1353/pbm.2005.0092>
  137. Peto, R., Doll, R., Buckley, J.D., and Sporn, M.B., Can dietary betacarotene materially reduce human cancer rates? *Nature*, 1981, vol. 290, no. 5803, pp. 201–208. <https://doi.org/10.1038/290201a0>
  138. *Users' Guides to the Medical Literature: Essentials of Evidence-Based Clinical Practice*, Guyatt, G., Meade, M.O., Rennie, D., Lee, P.R., and Cook, D.J., Eds., New York: McGraw Hill, 2008, 2nd ed.
  139. Stampfer, M.J. and Colditz, G.A., Estrogen replacement therapy and coronary heart disease: a quantitative assessment of the epidemiologic evidence, *Prev. Med.*, 1991, vol. 20, no. 1, pp. 47–63. [https://doi.org/10.1016/0091-7435\(91\)90006-p](https://doi.org/10.1016/0091-7435(91)90006-p)
  140. Petitti, D.B., Perlman, J.A., and Sidney, S., Postmenopausal estrogen use and heart disease, *N. Engl. J. Med.*, 1986, vol. 315, no. 2, pp. 131–132. <https://doi.org/10.1056/NEJM198607103150213>
  141. Petitti, D.B., Perlman, J.A., and Sidney, S., Noncontraceptive estrogens and mortality: long-term follow-up of women in the Walnut Creek Study, *Obstet. Gynecol.*, 1987, vol. 70, no. 3, pp. 289–293.
  142. Gage, S.H., Munafò, M.R., and Davey Smith, G., Causal inference in developmental origins of health and disease (DOHaD) research, *Annu. Rev. Psychol.*, 2016, vol. 67, pp. 567–585. <https://doi.org/10.1146/annurev-psych-122414-033352>
  143. Lobo, R.A., Hormone-replacement therapy: current thinking, *Nat. Rev. Endocrinol.*, 2017, vol. 13, no. 4, pp. 220–231. <https://doi.org/10.1038/nrendo.2016.164>
  144. Rossouw, J.E., Anderson, G.L., Prentice, R.L., et al., Risks and benefits of estrogen plus progestin in healthy postmenopausal women: principal results From the Women's Health Initiative randomized controlled trial, *J.A.M.A.*, 2002, vol. 288, no. 3, pp. 321–333. <https://doi.org/10.1001/jama.288.3.321>
  145. Dalton, S.O., Boesen, E.H., Ross, L., et al., Mind and cancer. Do psychological factors cause cancer? *Eur. J. Cancer*, 2002, vol. 38, no. 10, pp. 1313–1323. [https://doi.org/10.1016/S0959-8049\(02\)00099-0](https://doi.org/10.1016/S0959-8049(02)00099-0)

146. Johansen, C., Mind as a risk factor for cancer—some comments, *Psychooncology*, 2012, vol. 21, no. 9, pp. 922–926.  
<https://doi.org/10.1002/pon.3143>
147. Heikkilä, K., Nyberg, S.T., Theorell, T., et al., Work stress and risk of cancer: meta-analysis of 5700 incident cancer events in 116,000 European men and women, *Br. Med. J.*, 2013, vol. 346, art. ID f165.  
<https://doi.org/10.1136/bmj.f165>
148. Psychological stress and cancer, National Cancer Institute at the National Institute of Health, U.S. Department of Health and Human Services, 2012.  
<https://www.cancer.gov/about-cancer/coping/feelings/stress-fact-sheet>. Cited July 4, 2019.
149. Does stress affect cancer risk? Together we will beat cancer, Cancer Research, 2018. <http://www.cancerresearchuk.org/about-cancer/causes-of-cancer/cancer-controversies/stress>. Cited July 4, 2019.
150. Antonova, L., Aronson, K., and Mueller, C.R., Stress and breast cancer: from epidemiology to molecular biology, *Breast Cancer Res.*, 2011, vol. 13, no. 2, art. ID 208.  
<https://doi.org/10.1186/bcr2836>
151. Hjerl, K., Andersen, E.W., Keiding, N., et al., Breast cancer risk among women with psychiatric admission with affective or neurotic disorders: a nationwide cohort study in Denmark, *Br. J. Cancer*, 1999, vol. 81, no. 5, pp. 907–911.  
<https://doi.org/10.1038/sj.bjc.6690785>
152. Reiche, E.M., Nunes, S.O., and Morimoto, H.K., Stress, depression, the immune system, and cancer, *Lancet Oncol.*, 2004, vol. 5, no. 10, pp. 617–625.  
[https://doi.org/10.1016/S1470-2045\(04\)01597-9](https://doi.org/10.1016/S1470-2045(04)01597-9)
153. Soung, N.K. and Kim, B.Y., Psychological stress and cancer, *J. Anal. Sci. Technol.*, 2015, vol. 6, no. 30.  
<https://doi.org/10.1186/s40543-015-0070-5>
154. Koterov, A.N. and Biryukov, A.P., The possibility of determining of anomalies and pathologies in the offspring of liquidators of Chernobyl accident by the non-radiation factors, *Int. J. Low Radiat. (Paris)*, 2011, vol. 8, no. 4, pp. 256–312.  
<https://doi.org/10.1504/IJLR.2011.046529>
155. Antoni, M.H., Lutgendorf, S.K., Cole, S.W., et al., The influence of bio-behavioral factors on tumour biology: pathways and mechanisms, *Nat. Rev. Cancer*, 2006, vol. 6, no. 3, pp. 240–248.  
<https://doi.org/10.1038/nrc1820>
156. Li, P., Huang, J., Wu, H., et al., Impact of lifestyle and psychological stress on the development of early onset breast cancer, *Medicine (Baltimore)*, 2016, vol. 95, no. 50, p. e5529.  
<https://doi.org/10.1097/MD.0000000000005529>
157. Ross, L., Boesen, E.H., Dalton, S.O., and Johansen, C., Mind and cancer: does psychosocial intervention improve survival and psychological well-being? *Eur. J. Cancer*, 2002, vol. 38, no. 11, pp. 1447–1357.  
[https://doi.org/10.1016/s0959-8049\(02\)00126-0](https://doi.org/10.1016/s0959-8049(02)00126-0)
158. *Internal Report 14/002 of the Advisory Group to Recommend Priorities for IARC Monographs during 2015–2019, Updated 2018*, Lyon: Int. Agency Res. Cancer, 2014.
159. Gøtzsche, P.C., *Deadly Medicines and Organised Crime: How Big Pharma Has Corrupted Healthcare*, Boca Raton, FL: CRC Press, 2013.
160. Davey Smith, G., Data dredging, bias, or confounding. They can all get you into the BMJ and the Friday papers, *Br. Med. J.*, 2002, vol. 325, no. 7378, pp. 1437–1438.  
<https://doi.org/10.1136/bmj.325.7378.1437>
161. Feinstein, A.R., Clinical biostatistics. XLVII. Scientific standards vs. statistical associations and biologic logic in the analysis of causation, *Clin. Pharmacol. Ther.*, 1979, vol. 25, no. 4, pp. 481–492.  
<https://doi.org/10.1002/cpt1979254481>
162. Knottnerus, J.A., Alvan R. Feinstein, *J. Epidemiol. Commun. Health*, 2002, vol. 56, no. 5, pp. 322.  
<https://doi.org/10.1136/jech.56.5.322>
163. Feinstein, A.R., Meta-analysis: statistical alchemy for the 21st century, *J. Clin. Epidemiol.*, 1995, vol. 48, no. 1, pp. 71–79.  
[https://doi.org/10.1016/0895-4356\(94\)00110-C](https://doi.org/10.1016/0895-4356(94)00110-C)
164. *A Framework for Assessing Health Risks of Environmental Exposures to Children, EPA/600/R-05/093F*, Washington, DC: US Environ. Prot. Agency, 2006.
165. *IARC Monographs on the Evaluation of Carcinogenic Risks to Humans, Suppl. 7: Overall Evaluations of Carcinogenicity: An Updating of IARC Monographs*, Lyon, 1987, vols. 1–72.
166. *Radiation, Vol. 100D: A Review of Human Carcinogens*, Lyon: Int. Agency Res. Cancer, 2012.
167. Rothman, K.J. and Poole, C., A strengthening programme for weak associations, *Int. J. Epidemiol.*, 1988, vol. 17, no. 4, pp. 955–959.
168. Greenland, S. and Robins, J.M., Identifiability, exchangeability, and epidemiological confounding, *Epidemiol. Perspect. Innovations*, 2009, vol. 6, no. 4.  
<https://doi.org/10.1186/1742-5573-6-4>
169. Frick U. and Rehm J., Can we establish causality with statistical analyses? The example of epidemiology, in *Statistics and Causality: Methods for Applied Empirical Research*, Wiedermann, W. and von Eye, A., Eds., Hoboken, NJ: Wiley, 2016, pp. 407–431.
170. Andersen, H., History and philosophy of modern epidemiology, 2007. <http://philsci-archival.pitt.edu/4159/>. Cited July 4, 2019.
171. Charlton, B.G., Attribution of causation in epidemiology: chain or mosaic? *J. Clin. Epidemiol.*, 1996, vol. 49, no. 1, pp. 105–107.  
[https://doi.org/10.1016/0895-4356\(95\)00030-5](https://doi.org/10.1016/0895-4356(95)00030-5)
172. Charlton, B.G., The Zombie science of evidence-based medicine: a personal retrospective. A commentary on Djulbegovic, B., Guyatt, G.H. and Ashcroft, R.E., *Cancer Control*, 2009, vol. 16, pp. 158–168; *J. Eval. Clin. Pract.*, 2009, vol. 15, no. 6, pp. 930–934.  
<https://doi.org/10.1111/j.1365-2753.2009.01267.x>
173. Charlton, B.G., Are you an honest scientist? Truthfulness in science should be an iron law, not a vague aspiration, *Med. Hypotheses*, 2009, vol. 73, no. 5, pp. 633–635.  
<https://doi.org/10.1016/j.mehy.2009.05.009>
174. Charlton, B.G., The cancer of bureaucracy: how it will destroy science, medicine, education; and eventually everything else, *Med. Hypotheses*, 2010, vol. 74, no. 6,

- pp. 961–965.  
<https://doi.org/10.1016/j.mehy.2009.11.038>
175. Rose, S.P.R., *Molecules and Minds: Essays on Biology and the Social Order*, Chichester: Wiley, 1987.
  176. Charlton, B.G., A critique of biological psychiatry, *Psycho1. Med.*, 1990, vol. 20, no. 1, pp. 3–6.
  177. Fujimara, J.H. and Chou, D.Y., Dissent in science: styles of scientific practice and the controversy over the cause of AIDS, *Soc. Sci. Med.*, 1994, vol. 38, no. 8, pp. 1017–1036.
  178. Parascandola, M. and Weed, D.L., Causation in epidemiology, *J. Epidemiol. Comm. Health*, 2001, vol. 55, no. 12, pp. 905–912.  
<https://doi.org/10.1136/jech.55.12.905>
  179. Arnett, J.C., *The EPA's Fine Particulate Matter (PM 2.5) Standards, Lung Disease, and Mortality: A Failure in Epidemiology*, Washington, DC: Competitive Enterprise Inst., 2006.
  180. Bluming, A.Z. and Tavris, C., Hormone replacement therapy: real concerns and false alarms, *Cancer J.*, 2009, vol. 15, no. 2, pp. 93–104.  
<https://doi.org/10.1097/PPO.0b013e31819e332a>;  
 Bluming, A., Hormone replacement therapy. Judgment Day. What should we know? *Proc. Int. Symp. on Understanding Health Benefits and Risks, May 28*, Baltimore, MD: Johns Hopkins Univ., 2009.  
<https://jscholarship.library.jhu.edu/handle/1774.2/33639>. Cited July 4, 2019.
  181. Morabia, A., Hume, Mill, Hill, and the sui generis epidemiologic approach to causal inference, *Am. J. Epidemiol.*, 2013, vol. 178, no. 10, pp. 1526–1532.  
<https://doi.org/10.1093/aje/kwt223>
  182. Peeters, P.J.H.L., Etiological studies in complex diseases, *PhD Thesis*, Utrecht: Utrecht Inst. Pharm. Sci., 2016.
  183. Goodman, K.J. and Phillips, C.V., Hill's criteria of causation, in *Encyclopedia of Statistics in Behavioral Science*, Everitt, B.S. and Howell, D.C., Eds., Chichester: Wiley, 2005, vol. 2, pp. 818–820.
  184. Phillips, C.V. Gateway effects: why the cited evidence does not support their existence for low-risk tobacco products (and what evidence would), *Int. J. Environ. Res. Publ. Health*, 2015, vol. 12, no. 5, pp. 5439–5464.  
<https://doi.org/10.3390/ijerph120505439>
  185. Carl V Phillips, Tobacco Tactics. [https://www.tobaccotactics.org/index.php/Carl\\_V\\_Phillips](https://www.tobaccotactics.org/index.php/Carl_V_Phillips). Cited July 4, 2019.
  186. Gori, G.B., Epidemiologic evidence in public and legal policy: reality or metaphor? In *Critical Legal Issues: Working Paper Series No. 124*, Washington, DC: Washington Legal Found., 2004. <https://s3.us-east-2.amazonaws.com/washlegal-uploads/upload/GoriWP.pdf>. Cited July 8, 2019.
  187. Greenland, S. and Brumback, B., An overview of relations among causal modeling methods, *Int. J. Epidemiol.*, 2002, vol. 31, no. 5, pp. 1030–1037.  
<https://doi.org/10.1093/ije/31.5.1030>
  188. Phillips, C.V., Quantifying and reporting uncertainty from systematic errors, *Epidemiology*, 2003, vol. 14, no. 4, pp. 459–466.  
<https://doi.org/10.1097/01.ede.0000072106.65262.ae>
  189. Phillips, C.V. and LaPole, L.M., Quantifying errors without random sampling, *BMC Med. Res. Methodol.*, 2003, vol. 3, no. 9.  
<https://doi.org/10.1186/1471-2288-3-9>
  190. Phillips, C.V. and Goodman, K.J., The missed lessons of Sir Austin Bradford Hill, *Epidemiol. Perspect. Innovations*, 2004, vol. 1, no. 1, p. 3.  
<https://doi.org/10.1186/1742-5573-1-3>
  191. Terris, M., The Society for Epidemiologic Research and the future of epidemiology, *J. Publ. Health Policy*, 1993, vol. 14, no. 2, pp. 137–148.  
<https://doi.org/10.2307/3342960>

*Translated by M. Batrukova*