BRILL
WAGENINGEN
ACADEMIC

BENEFICIAL MICROBES (2026) DOI:10.1163/18762891-BJA00110

Beneficial Microbes
brill.com/bm

REVIEW ARTICLE

A pragmatic approach to integrate evidence-based medicine and personalized medicine: the example of personalized microbiome-targeting interventions

D. Zeilstra[1,2], A.A. te Velde[3], G. Remmers[4], I. Besseling-van der Vaart[5], R.J. Brummer[6] and A.D. Kraneveld[7]*

[1]Nutriz, Deventer, the Netherlands; [2]Microbiome Center, Zoetermeer, the Netherlands; [3]Amsterdam UMC location University of Amsterdam, Tytgat Institute for Liver and Intestinal Research, AGEM, Amsterdam, the Netherlands; [4]Stichting Mijn Data Onze Gezondheid (Foundation My Data Our Health), Amsterdam, the Netherlands; [5]Winclove Probiotics B.V., Amsterdam, the Netherlands; [6]Nutrition-Gut-Brain Interactions Research Centre, Faculty of Health and Medicine, School of Medical Sciences, Food and Health Center, Örebro University, Örebro, Sweden; [7]Division of Pharmacology, Utrecht Institute for Pharmaceutical Sciences, Faculty of Science, Utrecht University, Utrecht, the Netherlands; *dennis@microbiome-center.nl

Received 30 May 2025 | Accepted 5 January 2026 | Published online 21 January 2026

Abstract

Health care practitioners (HCPs) strive to provide the best medical care for each individual patient. The question as to what constitutes 'the best' does, however, not have a single straightforward answer. Evidence-based Medicine (EBM) and Personalized Medicine (PM) are two paradigms that have emerged as means to improve intervention selection. Both paradigms have their own strengths and weaknesses that affect their use in clinical decision-making. In this review we discuss the strengths and weaknesses from the patient's and HCP perspective: how to find the best intervention for a particular patient. We review methodological and practical aspects, and zoom out from the scientific level to the epistemological level to integrate EBM and PM. Both EBM and PM are based on a realist worldview and by adopting a pragmatist worldview the strengths of both paradigms can be combined. We apply this pragmatic approach, called Evidence-based Personalized Medicine (EBPM), to microbiome-targeting interventions. The example EBPM implementation uses four steps. First, it allows HCPs to provide information (clinical diagnosis, complaints, patient needs, laboratory measures) about an individual patient. Second, it uses a GRADE-based system to grade evidence of specific intervention components. Next, it combines the patient profile data and preferences with the graded evidence, to come to a suggestion for a personalized intervention. Finally, this method enables gathering of treatment effects providing feedback into the system and further improve suggestions for future patients.

Keywords

evidence-based medicine – personalized medicine – microbiome – epistemology

1 Introduction

Health care practitioners (HCPs) strive to provide the best medical care for each individual patient. The question as to what constitutes 'the best' does, however, not have a single straightforward answer. Careful appraisal of all available evidence often plays an important role, but various conceptual and epistemological perspectives exist on what contributes to evidence.

A well-known perspective is evidence-based medicine (EBM). In practice, the term EBM does not mean the same for every person. In protocols, research, and clinical practice, EBM often is used exclusively for medical practices (often for interventions) for which high quality scientific evidence is available. However, one of the original definitions given by Sackett *et al.* in 1996 describes that EBM 'means integrating individual clinical expertise with the best available external clinical evidence from systematic research' (Sackett *et al.*, 1996). While the Users' Guides about EBM published by JAMA since 1992 do put emphasis on critical appraisal of published research, these guides too note that 'many aspects of clinical practice cannot, or will not, ever be adequately tested. Clinical experience and its lessons are particularly important in these situations' (Guyatt *et al.*, 2000). In addition to evidence from systematic research and clinical expertise, later definitions added as a third pillar the role of the patients' preferences in selecting the best intervention (Haynes *et al.*, 2002). Together, these three pillars form the framework upon which EBM is based (Figure 1). The EBM perspective later expanded to disciplines other than clinicians and was referred to as evidence-based practice (EBP). EBP guidelines have been developed for a variety of HCPs, including nurses, physical therapists, and dieticians (Briggs Early and Stanley, 2018; Hillegass *et al.*, 2022; Mackey and Bassendowski, 2017). In the remainder of this article, we will use the more well-known term EBM but this is to be taken to involve EBP too and can be defined as: 'a problem-solving approach to clinical decision making that incorporates the use of best available evidence from systematic research, clinical expertise and assessment, and patient preferences and values' (Figure 1).

In recent years, personalized medicine (PM) has been embraced by many researchers and clinicians as an alternative paradigm to select the best possible intervention for a particular patient. Rather than merely using average effects found in clinical trials for a certain patient population (e.g. patients with a certain disease), the PM approach attempts to employ knowledge about specific patient signatures (e.g. their genome) as

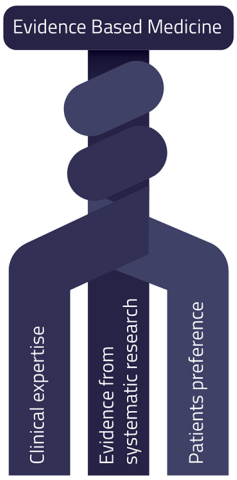

FIGURE 1 The three pillars of evidence-based medicine: evidence from systematic research, clinical expertise, and patient preferences. These three pillars are intertwined since each of the pillars can inform the other two.

means to more specifically select the best intervention for a particular patient. Although no commonly agreed definition exists, a useful definition is provided by the EU Health Ministers in their Council conclusions on personalized medicine for patients, which states that PM 'refers to a medical model using characterization of individuals' phenotypes and genotypes (e.g. molecular profiling, medical imaging, lifestyle data) for tailoring the right therapeutic strategy for the right person at the right time, and/or to determine the predisposition to disease and/or to deliver timely and targeted prevention' (Council conclusions on personalised medicine for patients, 2015). This field has broadened to different disciplines as well, including e.g. personalized nutrition (PN) (Ordovas *et al.*, 2018). For sake of clarity, we will use the term PM in the remainder of this text, but this may be taken to encompass specific disciplines such as PN as well.

In this article, we will assess the different approaches to integrate evidence into the daily practice from the perspective of the HCP and patient. We investigate how the different paradigms contribute to the selection of the best possible intervention for an individual patient, and we will show that this is a strikingly different goal than gaining knowledge about an intervention as such. As we will discuss, the EBM and PM paradigms seem to be at odds with each other, but by critically considering the epistemology of both paradigms, it is possible to integrate them. We will first briefly discuss the three pillars of EBM and the foundation of PM and put this into perspective with other factors that play a role in clinical decision-making. Next, we will zoom out from

the scientific perspective to an epistemological perspective to find a means to integrate EBM and PM in a way that suits the clinical practice. We will show that the gut microbiome is an interesting target for the prevention, treatment, and management of many diseases and that, due to the highly individualized character of the microbiome, an approach that integrates EBM and PM is preferable. Finally, we introduce a method for such integration of EBM and PM that has been developed and is used for microbiome-targeting treatments.

Although we will discuss the theoretic frameworks behind EBM and PM in detail, our focus is on the use of these concepts in the practice of patient care as caried out by HCPs, which is always directed at an individual patient. To make the conceptual reasoning more tangible, we will illustrate the various aspects by means of a hypothetical patient case (inspired by a real-life example). The example patient case that we will use is a woman, 60 years of age, who is obese (BMI 36) and has a fasting glucose of 8.2 mmol/l and HbA1c of 45 mmol/mol. She has hypertension, kidney stones, frequent urinary tract infections, and is often fatigued. Moreover, she suffers from anxiety and depression, and has many gastrointestinal (GI) complaints such as abdominal pain, varying defecation frequency and stool consistency, bloating, belching, fatty stools, and food intolerances.

2 EBM's 1st pillar: scientific clinical evidence

Randomised controlled trials

Scientific clinical evidence in the context of EBM is often based on randomized controlled trials (RCTs) because these are considered the 'gold standard' in the so-called 'evidence pyramid' (with meta-analyses of RCTs being considered the only type of study exceeding the strength of this 'gold standard') (Burns *et al.*, 2011).

The very reason that RCTs are considered to provide evidence with the highest degree of certainty is that, when properly designed and conducted, their design eliminates most sources of bias (Pan *et al.*, 2018). RCTs have a so called high internal validity (Jüni *et al.*, 2001). Internal validity is defined as the 'extent to which systematic error (bias) is minimized in clinical trials' (Jüni *et al.*, 2001). In other words, conclusions that are drawn within the context of such study are very likely to be valid. Other study designs, such as case-control or observational studies are hampered by confounding factors that cannot be excluded or corrected for completely, and the degree of certainty that an observed effect is due to the intervention is therefore lower (Pan *et al.*, 2018). This means that the RCT approach is the best available method to systematically investigate the efficacy of an intervention in a specific homogeneous population. Efficacy, however, is not the same as effectiveness. Efficacy is defined as the extent to which an intervention produces its intended effect under ideal circumstances, such as in a RCT, while effectiveness is defined as how well an intervention works under 'real life' conditions (Ernst and Pittler, 2006; Marley, 2000). In other words, if we want to know the average effect and effect size of an intervention as such (i.e. its efficacy), excluding to the best of our abilities the influence of intervention-independent factors that may affect the intervention outcomes, the RCT methodology provides the technique to study this. In real life situations a large range of factors may affect this outcome, leading to different effectiveness.

One of the reasons that RCTs have a high internal validity is that the intervention and control (often placebo) groups are made to be exchangeable (Hernán and Robins, 2022). In other words, it should not matter which of these groups gets the placebo intervention and which receives the *verum* intervention, as one may expect both groups to respond the same to the same intervention. Importantly, this can be quite challenging to achieve in practice. An important way in which this exchangeability is approached as closely as possible, it by using clear-cut inclusion and exclusion criteria and a sound randomization, stratification, and allocation procedure to realize a homogeneous study population. This leads to a study population that can be rather specific, and which is not necessarily representative for the patient population seen by clinicians. As we will show later, it is very unlikely that our example patient would be part of such a study population due to her rather complex, multimorbid profile.

Summarizing, the RCT approach is the best available method to systematically investigate the efficacy of an intervention, but this often investigated in a very specific homogeneous study population that may not translate easily to real-world patients. This is the Achilles' heel of the RCT method and the way it is often employed: the external validity is often poor.

A limitation of the RCT method: poor external validity

Suppose that a study includes only Caucasian males aged 20-30 who have a single health problem (which is the focus of the study) and who otherwise are healthy. What does the result of that study mean for a 60-year-old Asian woman with multiple comorbidities, like our

example patient? In other words, due to strict inclusion and exclusion criteria, the generalizability of the evidence from RCTs may be questionable and the study population may not be representative for the patient sitting in front of the practitioner (Dekkers *et al.*, 2010). Social-psychologic factors that are not linked to the RCT method as such, can further exacerbate this narrowing of the study population by the way the RCT method may be employed in reality. These social-psychologic factors include e.g. publication pressure (negative trials are much less often submitted and accepted for publication (Hopewell *et al.*, 2009)), desire to publish in high-impact journals (for small trials the most prestigious journals seem to have a bias towards favourable effects (Siontis *et al.*, 2011)), or financial interests associated with positive results (Lundh *et al.*, 2017). Whatever the reason, due to this attitude researchers want to make the chance of success as high as possible and tend to include only those patients in which the largest effects are expected, potentially making the study population less representative for the real-world population. On the one hand a narrowly defined study population helps to identify real-world patients for whom the results of a specific RCT may apply, on the other hand it makes it less certain to what extent the effects found in such trial will reproduce in patients that differ from the defined study population.

Several reviews have assessed the representativeness of actual RCTs. For example, a meta-review of 52 studies that investigated the representativeness of RCT populations compared to real-world patients, conclude that over 70% of the studies explicitly concluded that RCT samples were not broadly representative of real-world patients (Kennedy-Martin *et al.*, 2015). The authors found e.g. that reviews that assessed cardiology and oncology RCTs found that real-world patients were often older than those included in the RCTs (Kennedy-Martin *et al.*, 2015). Another systematic study of over 43,000 RCTs found that over 90% excluded multimorbidity, while the prevalence of multi-morbidity in real-world patients is 41% (Tan *et al.*, 2022).

Summarizing, the RCT method is the best means to exclude most sources of bias, therewith providing high internal validity, but may result in a poor external validity. This is important to take into account when assessing the value of evidence from the HCP's and patient's perspective, as they are interested in the potential effects of an intervention for that particular patient. Our example patient and her HCP want to find the intervention that has the best chance of success for her specifically.

The role of causal thinking in the interpretation of evidence from RCTs

Another important aspect to take into account is that RCTs attempt to assess the efficacy of an intervention, with the effect size being the *average* difference in effect between the placebo and verum intervention groups (Hernán and Robins, 2022). In essence, this approach assigns the effect to the intervention as if it were a property of the intervention. An example of this is the commonly used phrase 'paracetamol cures headache'. Contrary to what this phrase expresses, reality is that paracetamol may improve headache symptoms in person A, but not in person B (even with the same type of headache). This in itself shows that 'cures headache' is not a property of paracetamol as such and depends on the characteristics of a person. Nevertheless, it is understandable that effects of interventions are often considered to be property of that intervention, given the dominant role in society (and science) of classic cause-effect conceptual thinking that we know from everyday physics (dropping a stone in water *causes* waves). This rather mechanistic and reductionistic world view has become so dominant because its huge success in physics marked the start of modernity and modern science (Toulmin, 1990). Causal thinking not only obfuscates interpretation of intervention studies, it has in fact played a leading role in epidemiology as a practical discipline concerned with identifying causes of disease, even though in many diseases a multitude of interrelating factors are known to play a role and the classic cause-effect concept is not valid (Galea *et al.*, 2010; Sturmberg and Marcum, 2023). In addition to this fundamental problem of causal thinking, the practical application of causal inference in epidemiology is often problematic, because various forms of bias and confounding (e.g. reverse causation or e.g. the Simpson's paradox) can blur causal relationships (Hernán and Robins, 2022; Julious and Mullee, 1994).

Back to the relevance for EBM, for most interventions a study outcome is in fact a probability of an effect in comparison to the placebo treatment (paracetamol completely resolves headache in A percent of the study population, partially in B percent, doesn't help in C percent, and worsens it in D percent). Although in many studies this kind of information is not available, the key issue remains the same: within the study population the intervention effect varies between the individuals. Except for rare cases, often it is simply not known which individual does respond and who does not, even in a narrowly-defined study population (Kelley and Kaptchuk, 2010; Kravitz *et al.*, 2004). The fact that

many people that receive an intervention do not experience substantial benefit from it is also illustrated by sometimes high numbers needed to treat. The number-needed-to-treat expresses how many patients have to be treated in order to benefit one, and this can be as high as 446 for e.g. statins for prevention of major vascular events in people over 75 without vascular disease (Heneghan and Mahtani, 2019).

Summarizing, it can be concluded that even within an apparently homogeneous study population considerable variety in intervention-response continues to exist. As such, no RCT can predict the response of an individual patient sitting in front of the HCP. This means that even in the unlikely case that there would exist a RCT of an intervention in a population with the exact profile of our example patient, her HCP and she would still not be able to tell if she will benefit from that intervention. At best, the RCT provides a probability of an effect. If we can resist the inclination of causal thinking and use a probabilistic perspective instead, evidence from RCTs can help to find the intervention with the best probability of success.

3 EBM's 2nd pillar: individual clinical expertise

Translation to multimorbidity of studies involving one health problem

The vast majority of patients who require health care are burdened with chronic diseases. Chronic diseases are typically non-communicable, usually of long duration, progresses slowly, and are often the result of a combination of genetic, physiological, environmental and behavioural factors (World Health Organization (WHO), 2022). Moreover, the vast majority of patients with chronic diseases have a multimorbid profile. In fact, multimorbidity is 'the most common chronic condition' (Tinetti *et al.*, 2012). Multimorbidity is 'the coexistence of multiple chronic diseases or conditions' (Tinetti *et al.*, 2012) and is typically characterized by systemic low-grade inflammation, insulin resistance, disturbed HPA-axis, and similar dysregulation of systemic control mechanisms. Our hypothetical patient provides a typical example of multimorbidity, with various health conditions (e.g. hypertension, kidney stones, urinary tract infections, fatigue, anxiety, depression, and GI complaints), as well as systemic dysregulation exemplified by her fasting glucose that indicates insulin resistance. The profile of health problems and complaints is typically unique for a patient, but, as explained later, often the underlying disturbed pathways are more general.

Although multimorbidity is common, it is rarely the basis for studies investigating interventions, since most clinical studies involve patients with one specific condition and often explicitly have multimorbidity as an exclusion criterion. In multimorbid patients, evidence from studies with specific interventions for specific health problems is only partially informative. That is, these kinds of studies do neither provide an answer to the question whether or not multimorbid patients will benefit from a specific intervention, nor what intervention to use for each of the various health problems of the patient. For example, there may be high quality evidence for intervention of type 2 diabetes (T2DM) and other high-quality evidence for intervention of major depression disorder (MDD). However, even though it is known that 1 in 4 patients with T2DM suffers from clinical depression (Hamer *et al.*, 2019), no clinical studies can be found of an intervention that treats both health problems. This means that the HCP treating our example patient (who has both signs of diabetes as well as depression) cannot solely rely on evidence from RCTs to find a proper treatment for her.

An aspect linked to multimorbidity is the concept of the life course perspective. This concept entails the notion that a patient's health state at any point in time is also influenced by the preceding life course and, conversely, that influencing one's life course at an early stage of disturbance can prevent or substantially delay development of diseases and/or comorbidities (Blane *et al.*, 2015). For example, T2DM typically develops over the course of a decade or more. In addition, the life course perspective acknowledges that social and personal circumstances play a significant role in one's health. For example, it is known that depression affects one's partner's cognitive functioning (Blane *et al.*, 2015), and it may be necessary to take into account these personal circumstances in the treatment of patients. These highly individual aspects cannot be encompassed in any systematic approach or protocol, and this is where clinical expertise plays an important role.

One approach to deal with multimorbidity is to investigate whether there is a common underlying pathological mechanism (Sturmberg *et al.*, 2017). Typically, the route to advance knowledge about mechanisms underlying comorbid health problems is to conduct mechanistic studies to identify commons mechanisms (Mode of Action studies), translate this into potential interventions targeting comorbidities simultaneously, validate this in interventions studies (initially Proof of Concept studies), and update protocols for HCPs. This route could work, but as previously mentioned, is also

strongly limited due to problems with generalizability and occurs only to a limited extent due to the preference of trials with a single health problem in a homogenous population.

Using the aforementioned example of MDD and T2DM, it is known that insulin resistance and systemic low-grade inflammation (two core mechanisms underlying T2DM) are mechanisms that are also underlying to MDD (Chan *et al.*, 2019; Lyra e Silva *et al.*, 2019). To translate this into clinical practice, insulin resistance can typically be identified using a HOMA-IR panel, elevated HbA1C or fasting glucose. In our example patient, the latter two markers were elevated. Moreover, the typical signs of metabolic syndrome (including hypertension and overweight that were present in our example patient) can also indicate insulin resistance. Systemic low-grade inflammation is characterized by elevated pro-inflammatory cytokines (e.g. IFN-γ, IL-6, or TNF-α), or elevated CRP. As discussed later, we will use the concepts addressed in this manuscript in the context of gut microbiome dysbiosis. Since our example patient is inspired by an actual case and no blood levels were available, we can use biochemical markers in faecal analysis results as a proxy for systemic inflammation as well. This can include elevated faecal calprotectin, which was present in our example patient, which is known to correlate with systemic markers of inflammation (Bourgonje *et al.*, 2018; Heinzel *et al.*, 2024). Moreover, symptoms like fatigue (also present in our example patient) often accompany systemic low-grade inflammation and insulin resistance. It is known that these two mechanism can be targeted with metformin, the first-line intervention for T2DM (Park *et al.*, 2018), but only a single study on the use of metformin in MDD has been published (and has been retracted) (Abdallah *et al.*, 2020, 2022). This illustrates the limitations of the typical path from common mechanism discovery to intervention development and validation.

Summarizing, multimorbidity and life courses are highly individual aspects and the HCP's clinical expertise plays an important role in addressing these and taking them into account in selecting an intervention for a particular patient. Clinical expertise is not only formed through experience, but also via primary education, refresher trainings, or e.g. pharmacotherapeutic transmural consultations. In addition, training and use of diagnostic tools, and working with external labs that provide proper support are part of the skills of HCPs that form the core of the second pillar of EBM.

Translation of studies with one intervention in many patients to many intervention options for one patient

The first pillar of EBM is about gathering and interpreting the best available scientific evidence from systematic clinical study. It addresses what can be concluded about the efficacy of an intervention in a defined population. The second pillar of EBM is to connect this with the reality of the clinical practice. More precisely, this is about determining what is the best approach for an individual patient. This second pillar revolves around the 'individual clinical expertise', as the aforementioned definition of EBM calls it (Sackett *et al.*, 1996). In this second pillar, the input from the first pillar (scientific clinical evidence) is to be assessed and applied to an individual patient and this requires the experience, know-how, and skills of the HCP. As will be discussed later, the concept of 'real-world evidence' and phase IV studies can play a key role in connecting this second pillar back to the first, providing a cycle in which understanding and knowledge advances.

Although the perspective that considers effects as properties of an intervention is useful and necessary for companies and authorities, for HCPs the individual patient perspective is pivotal. In a sense, this perspective is the other way around (Figure 2). That is, HCPs are burdened with the task to determine what would benefit an individual patient, or, in other words, what the effect of an intervention can be in a particular patient. This is strikingly different from determining the efficacy of an intervention. While the latter is determined by administering one intervention to a large number of patients (i.e. in a RCT, Figure 2A), the HCPs task requires to weigh the potential efficacy of a number of interventions with respect to their one patient (Figure 2B). With our hypothetical patient as example, the HCP is burdened with the task to determine which intervention(s) may help to combat the various health problems, and for each potential intervention one or more efficacy studies may be available.

Given the fact that individuals can respond very different to the same intervention, even in a narrowly defined population (Kelley and Kaptchuk, 2010; Kravitz *et al.*, 2004), such investigation to find the best intervention for an individual is desired if it is attainable. However, to assess and try every potential intervention is of course practically, financially, medically, and ethically impossible. In practice, a limited number of different interventions is sometimes employed on an individual, but this merely approaches the aforementioned theoretical concept of systematically testing all potential interventions. Thus, some other means to determine

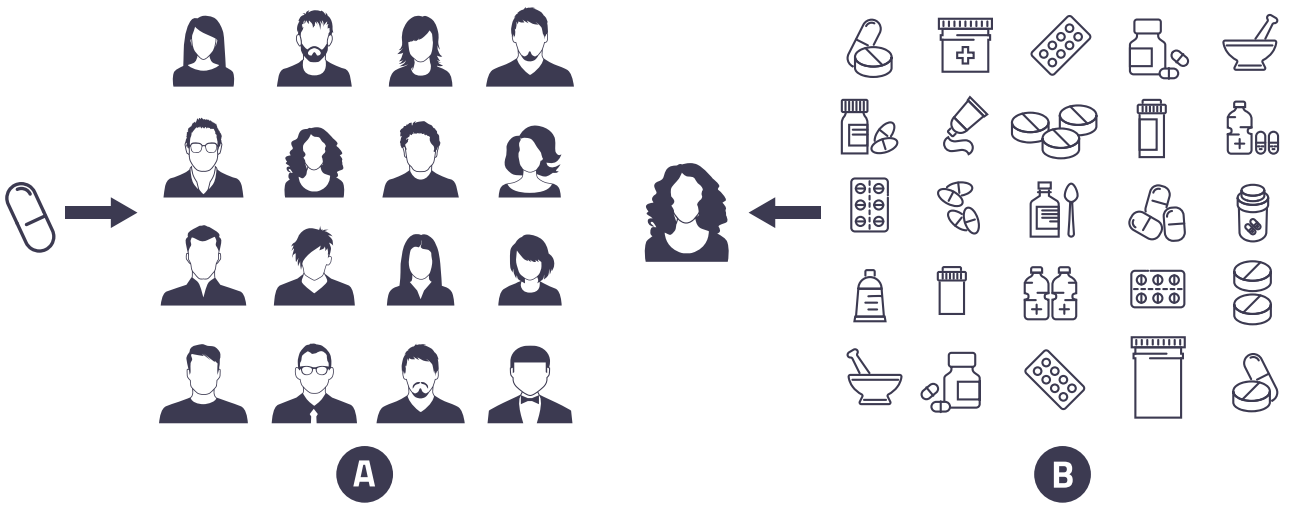

FIGURE 2 The different views of research and practice. (A) Researchers are faced with the challenge to determine the effect of a single intervention. To assess this, they administer that single intervention to multiple people and consider the average effect to be the efficacy. (B) In contrast, medical doctors are faced with the challenge to find from multiple intervention options the one with the best effect for a single patient. To assess this, they would ideally administer all intervention options to that single patient and select the intervention with the best effect.

what works best for the individual patient is needed and here individual clinical expertise comes in: the selection of specifically targeted intervention(s) is made using the HCP's clinical expertise.

Translation of mechanisms to the clinical practice

Alternatively, HCPs may attempt to identify common mechanisms of health problems of individual chronically ill patients themselves via literature review, in order to select the best intervention. Unfortunately, as shown by the aforementioned example, they are usually empty-handed with regard to evidence from clinical studies. Most intervention studies are simply not designed to answer this question, whereas the larger number of mechanistic studies that do aim to uncover mechanisms typically do not provide evidence about interventions and often lack translation to and validation of interventions. Exceptions are studies with nutraceuticals (e.g. vitamins) or, as discussed later, microbiome-targeting interventions such as prebiotics and probiotics. For these kinds of interventions often many different studies have been conducted, some assessing efficacy and other mechanisms of action. Hence, these kinds of interventions may offer more options for integrating mechanistic knowledge of health problems (i.e. mechanistic pathways, aetiology, and pathogenesis) and evidence about the mechanistic and clinical effects of these interventions.

In addition, the sheer number of publications makes it extremely challenging for HCPs to figure out by themselves which common mechanisms they may be dealing with for every individual case. Due to time constraints, one must often rely on guidelines published by their peers, or reviews and appraisal publications, which may introduce the same kind of potential bias that was reason to introduce EBM in the first place (Guyatt, 1992). This includes for example influence of personal views or not capturing the full width of the available research. Advances in the understanding of mechanisms sometimes become a target for treatment-validation in clinical trials. In addition, or alternatively, these insights find their way into courses, refresher trainings, or pharmacotherapeutic transmural consultations that HCPs follow to remain up to date. Via the latter and personal experience, this kind of knowledge becomes part of HCPs' personal clinical expertise. Moreover, new insights become incorporated in basic educational programs for HCPs, enabling younger generations of HCPs to employ this knowledge.

Thus, to identify and understand common underlying mechanisms for a particular patient, the HCP's individual clinical expertise is key and direct reliance on evidence plays a minor role. This route of growing clinical expertise occurs next to the experience that HCPs gain by treating their patients with interventions based on evidence from intervention studies and/or clinical protocols. Over time, HCPs learn in practice when an intervention is likely to be successful and what the effects on comorbidities can be.

4 EBM's 3rd pillar: patients' preferences

The first two pillars of EBM are about gaining and translating knowledge about potential interventions for specific health problems and using that to find the intervention that has the best chance to bring about the desired therapeutic effect in a particular patient. The third pillar is about the question how and by who that knowledge must be valued. That is, all interventions (including no intervention at all) have potential desired and undesired effects and the decision about which intervention may meet the needs of the patient the best, is about values (Kelly *et al.*, 2015). This includes valuing the chance to obtain a certain effect, as well as valuing which effects are most desirable and which side effects will have the biggest impact on the quality of life. For example, for a patient with an active sex life the risk of sexual dysfunction as a side effect of a certain intervention might outweigh the better chance of therapeutic effect of that intervention compared to alternatives (Kelly *et al.*, 2015). Here, multimorbidity also plays a role, as the treatment decision can involve which health problem to treat first or pay the most attention to. In these value-driven decisions, the preferences of the patient play an important role that cannot be captured by either evidence or clinical expertise. The same applies to social, cultural, or religious factors, which are also important factors that affect the patient preferences and thus the treatment decision. The patient preferences as third pillar of EBM are not a one-way-path, since information from the HCP may affect the patient preferences. Hence, it is crucial that the HCP conveys as much information as possible to the patient about the what, why, when, how of treatment options. In other words, informed consent should be considered part of the third EBM pillar.

Unfortunately, health care systems, policymakers, regulators, and insurance companies have become major factors in the decision-making process of HCPs, as they influence this via guidelines, protocols and reimbursement policies that do not necessarily take into account the patients' individual situation and preferences (Felder *et al.*, 2021). Similarly, while patients value health related quality of life and this has gained much attention in scientific research in the past two decades, this often does not translate into reimbursements of treatments that aim to improve quality of life (Eichelberger *et al.*, 2021). In fact, payers like insurance companies 'usually care most about clinical efficacy and cost' (Eichelberger *et al.*, 2021). Patients' preferences may differ from the clinicians' therapeutic goals or from policymaker goals. For example, a study among IBD patients showed that therapeutic goals proposed by physicians, such as healing the mucosal lesions, are not a priority for most patients (Casellas *et al.*, 2017).

One way forward is to improve HCP education, guidelines, and protocols to better represent patients' preferences (Montori *et al.*, 2013). This route is, however, paved with difficulties and cannot incorporate preferences of individuals. Another option, therefore, is to better integrate patients' preferences (e.g. input about which complaints have the biggest impact on their quality of life) in the clinical decision-making process. That is, explicitly incorporate discussion about this in the consultation and ensure that there is a shared decision about treatment.

5 Personalized Medicine

The second paradigm for the use of scientific evidence in clinical practice is PM. The PM paradigm originated in the beginning of this century, in the era of the Human Genome Project (Emmert-Streib, 2013). It is based on the idea that newly developed omics techniques provides better insights into the inter-individual variation in patient responses (McCoy, 2020). More recently, artificial intelligence and wearable smart sensors have been proposed as additional means to gain insights into individual responses (Lin and Wu, 2022). The premise of PM is that better understanding of factors that cause differences in intervention response combined with utilization of more measurements of patients can improve the prediction of the chance of success of an intervention for a particular patient and therewith aid the decision-making process.

One of the techniques employed within the PM paradigm is to use subgroup analysis to identify patient characteristics that may distinguish responders from non-responders (Kent *et al.*, 2018). As discussed before, EBM typically uses average outcomes from RCTs to determine the efficacy of an intervention, but several examples exist in which opposing effects occurred in subgroups within the study population that cancelled each other out, yielding an average null effect (Blackstone, 2019; McCoy, 2020). Despite these average null findings, the patients in the subgroups with a positive intervention response did in fact benefit from the intervention and PM is about identifying responder-discriminating factors. Since within the PM paradigm more clinical trial data is used to improve the generalizability of findings, this puts pres-

sure on data processing and interpretation capabilities. Artificial intelligence and machine learning are thought to be capable of improving the identification of responder-distinguishing factors (Lin and Wu, 2022; Subbiah, 2023). In microbiome studies, machine learning and deep learning models have already been successfully applied to gain deeper insights into specific features, such as phenotyping, spatial and temporal distribution, and prediction of host responses (Hernández Medina *et al.*, 2022). As an example, a supervised deep neural network that integrates gut metagenomics, plasma metabolomics, immune cell profiling, blood laboratory data and detailed clinical symptoms, was able to predict clinical severity and identify disease- and symptom-specific biomarkers in myalgic encephalomyelitis/chronic fatigue syndrome (Xiong *et al.*, 2025).

As mentioned in the introduction, PM has expanded into various disciplines, including nutrition. Personalized nutrition (PN), also referred to as 'precision nutrition', may be considered a specialization of PM and, since we focus in this article on microbiome-targeting interventions and many of these are often considered nutritional rather than medical interventions, the PN paradigm deserves attention here as well. For PN no clear definition exists either, although it is described the American Nutrition Association as 'a field that leverages human individuality to drive nutrition strategies that prevent, manage, and treat disease and optimize health' (Bush *et al.*, 2020). Although many of the more traditional studies investigating the link between diet and health (e.g. studies investigating the impact of saturated fat on health outcomes like coronary heart diseases) have not taken into account the gut microbiome at all, within the PN paradigm a prominent role of the gut microbiome has been identified as one of the factors that needs to be taken into account in the personalization of dietary interventions (Bush *et al.*, 2020; Ordovas *et al.*, 2018; Torres and Tovar, 2021).

Summarizing, within PM and PN, the evidence from systematic studies is focused more on external validity and translatability of findings to individual patients. In other words, dissecting study populations into smaller subgroups and potential introduction of bias are leveraged against gaining more information that can help to identify which patients may be responders. From the HCP and patient perspective, PM and PN mean that a larger amount of patient data is used to select the intervention with the best probability of success. As exemplified by the aforementioned study on myalgic encephalomyelitis/chronic fatigue syndrome (Xiong *et al.*, 2025), this larger amount of data includes the use of biomarkers as a means to categorize a patient into specific groups for which specific interventions are deemed more effective.

6 How to combine EBM and PM: assessing the underlying epistemology

The PM paradigm is often considered to be epistemologically opposed to EBM, as EBM is based on population-based or average effects, whereas PM is based on understanding of disease aspects unique to the individual (McCoy, 2020; Tonelli and Shirts, 2017). From the EBM perspective, the kinds of *post hoc* analyses conducted within the PM approach are often considered prone to bias or confounding (Senn, 2018). Subgroup analyses can be obfuscated by e.g. using arbitrary outcomes measures, or arbitrary thresholds of a continuous variable to define responders or non-responders. On the other hand, the aforementioned examples of null findings due to subgroups with responses that cancelled each other out, provide compelling examples that show how PM's subgroup analyses can identify groups of patients that can in fact benefit from the intervention.

To come to an approach that integrates the strengths from EBM and PM, it is not sufficient to assess scientific insights only. The fact that EBM and PM are seemingly at odds with each other lies not at scientific evidence level, but at epistemological level. That is, one may emphasize the potential bias introduced by subgroup analyses or the potential gain in responder identification, or one may emphasize the problems with translatability or the excellent internal validity of RCTs. Whatever emphasis one uses, this has nothing to do with evidence as such, but much more with ones presuppositions and worldview through which the evidence is interpreted (Gauch, 2003). Importantly, presuppositions by definition cannot be proven or disproven (Gauch, 2003; Zeilstra *et al.*, 2018), so disputes about interpretations of studies or methods that are caused by different presuppositions cannot be settled with more evidence. Therefore, it is required to zoom out and critically review the epistemological perspectives underlying the EBM and PM paradigms.

Interestingly, the PM and EBM perspectives share the premise that treatment decisions that are solely based on clinical experience and expert judgment are prone to all sorts of cognitive biases, and that using a more data-driven approach generally improves decision-making (Kent *et al.*, 2018). This common presupposition pro-

vides a first aspect that can be used to integrate the EBM and PM paradigms. In addition, it points out another epistemological overlap between PM and EBM, namely the presupposition that evidence from scientific research leads to generalizable conclusions that can aid the decision-making process. For EBM this means the presupposition that the average effect found in clinical trials is generalizable to a population with characteristics that are similar to the study population. For PM the presupposition is that data from clinical studies can be used to identify traits that predict a good response to a certain intervention and that this response is generalizable to a population with that trait.

In both cases this reveals a so called realist worldview, the philosophical perspective that assumes there is one reality that we can know via through systematic research (Chakravartty, 2017). In EBM and PM, this is reflected among others by the rather reductionistic view that assumes that by understanding the separate factors, we can understand the whole. For example, investigators tend to dissect things into ever more details, e.g. via 'omics' methods (e.g. genomics, epigenomics, transcriptomics, metabolomics, proteomics, etc.), assuming that knowledge of these aspects can help to predict individual responses (Bush *et al.*, 2020). Such reductionistic approach is highly effective in physics, because in many cases the superposition principle applies: the net response caused by multiple factors is the sum of the responses caused by each factor individually (Illingworth, 1991). However, superposition applies in linear systems, and predicting a response is much more difficult or fundamentally impossible if a system is either non-linear, or regulated, or interactive. In physiology all three apply, which makes it questionable if individual responses can be predicted via reductionistic studies. Importantly, this is not to say that omics methods will not be able to advance scientific understanding, which we believe it will. The point here is that these methods are in many cases (implicitly) used within a realist world view, which as such has limitations.

Secondly, it can be argued that studies of diseases start with the definition of the disease, which is at least in part a construct of the human mind rather than objective reality. For example, some scholars argue that attention-deficit hyperactivity disorder (ADHD) is a 'hypothetical construct' rather than a 'objective fact of nature' (Tait, 2005; Thurber *et al.*, 2009). Scientific findings such as differences in brain scans *start with the definition* to divide participants into cases and controls, rather than that ADHD objectively follows from the brain scans (Te Meerman *et al.*, 2020). Together, these examples show that at least part of our understanding of the world starts with our constructs rather than being merely a discovery of reality as is. An opposing philosophical worldview that acknowledges this is the constructivist view. This philosophical perspective states that every interpretation of the reality is in essence a construct of our mind, and influenced by the context, our social interactions, our past experiences, our perspective, and our expectations (Niiniluoto, 1991). Examples of optical illusions, such as the Ames room (Ames Room (Philip Zimbardo), 2010), as well as the aforementioned example of defining diseases such as ADHD show that at least to some extent our understanding and interpretation of the world indeed is based on a mental construct. However, although the constructivist worldview acknowledges that all interpretations and conclusions that we draw are at least in part influenced by our presuppositions, this worldview poses the risk of sceptically rejecting any form of evidence as useful.

With the HCPs' and patients' search for the best intervention for a specific patient by being a highly practical enquiry, and clinical work being an applied profession, a better suited philosophical perspective might be a third worldview: pragmatism. The pragmatist worldview accepts that there is a single world, but also that our understanding of it is always coloured by our presuppositions (Legg and Hookway, 2021). In essence, the pragmatist worldview accepts things to be true if they work.

The pragmatist worldview allows to use the best of both worlds of EBM and PM. For example, from a pragmatic point of view it can be accepted that some results from clinical trials are simply so compelling and studied in such a broad population, that it is reasonable to consider the average effect found in studies to be the best predictor for the effect in an individual patient. On the other hand, it can also be pragmatically accepted that some subgroup analyses provide such a compelling case that it is reasonable to consider that the subgroup-defining trait is a good predictor for the effect in an individual patient. Moreover, this view also allows consideration of mechanistic insights at different levels of detail. For example, in some cases a factor that can be predictive for an intervention effect may be found at molecular level, e.g. a genetic polymorphism causing an altered expression of some receptor. In other cases, such subgroup-defining trait may be identified at less detailed level, for example the aforementioned role of insulin resistance in T2DM and MDD (which can be due to a wide range of molecular mechanisms). In fact, in some cases there can be a trait that distinguishes

responders from non-responders for which the mechanism of action is simply unknown, yet if the evidence is compelling, a pragmatist worldview could consider it reasonable to use the trait in the decision-making process. The pragmatist worldview is also reflected in the notion that clinical practice informs evidence and this is why some authors have argued that Practice-Based Evidence (PBE) should be embraced rather than tolerated next to EBM (Green, 2008; Greenhalgh, 2020; Ogilvie *et al.*, 2020). Moreover, PBE enables a more flexible input of patients and citizens in research, and thus opens the door for Citizen Science (Guerrini *et al.*, 2022; Hsueh *et al.*, 2017).

Here, we will introduce a pragmatic approach that integrates EBM and PM, using gut microbiome-targeting interventions as an example. First, we will introduce the (gut) microbiome and explain why gut microbiome-targeted interventions are good examples for such an approach.

7 Microbiome and microbiome-targeted interventions

The human gut microbiome is defined as: all microorganisms (e.g. bacteria, fungi, viruses, protists, etc., together called microbiota) plus the 'theatre of activity' (e.g. signalling molecules, microbial structural elements, etc.) (Berg *et al.*, 2020). Research in the last decades has associated the gut microbiome [or to be more precise: the stool microbiome, as most studies use feacal analyses] with a large range of health issues, including gut-related health problems such as irritable bowel syndrome (IBS) and inflammatory bowel disease (IBD), as well as systemic health problems such as obesity, liver diseases, T2DM, cardiovascular diseases, allergies, cancer, MDD, autism, neurodegenerative diseases, and auto-immune diseases such as type 1 diabetes, multiple sclerosis, or rheumatoid arthritis (Christovich and Luo, 2022; Gomaa, 2020; Miyauchi *et al.*, 2023). Microbiome-targeted interventions can consist of diet, antibiotics, prebiotics, probiotics, postbiotics, or faecal microbiota transplantation (FMT). Several aspects make microbiome-targeted interventions a good example to pragmatically integrate EBM and PM.

First, microbiome-targeted interventions typically interact with the commensal microbiota. It is known that the gut microbiota is as unique to individuals as a fingerprint (Franzosa *et al.*, 2015; Tierney *et al.*, 2019), and has compositional changes over time, even daily, while also having a certain long term individual stability or resilience (Fassarella *et al.*, 2020; Revel-Muroz *et al.*, 2023; Vandeputte *et al.*, 2021). This makes it reasonable to expect that the same intervention will not induce the same effect in everyone and that a more personalized approach can be beneficial. Furthermore, some of the microbes administered as part of the intervention may multiply within an individually unique intestinal microbial ecosystem and the functionality of these 'offspring' will be affected by the conditions in this specific ecosystem (van Baarlen *et al.*, 2009; Zhang and Knight, 2023). Thus, a more personalized approach to gather and understand evidence about interventions seem to make sense.

Second, the number of microbiome-related factors that might distinguish between responders and non-responders is very high. As an example, while the human genome consists of some 63 thousand genes (Nurk *et al.*, 2022), the estimated number of microbial genes in our gut microbiome is over 22 million (Tierney *et al.*, 2019). Although it is unlikely that many of these genes will act as a single discriminating factor for responders, this large number of genes illustrates that identifying specific discriminators at gene or molecular level is challenging, let alone testing each potential discriminating molecular factor in a clinical trial. Here, a more pragmatic approach can be of help, for example by using higher level mechanistic insights about the role of the microbiome (i.e. not at molecular level), such as the aforementioned insulin resistance or systemic low-grade inflammation.

Third, previously we have demonstrated that RCTs with probiotics may not always provide valid results (Zeilstra *et al.*, 2018). We identified three presuppositions behind RCTs the must be true for the conclusions to be valid. The first presupposition is, as mentioned before, that the verum and placebo groups are exchangeable at group level. Secondly, it is presupposed that there are no uncontrolled factors that influence the outcome of interest, and, thirdly, it is presupposed that the intervention is well-defined. When these presuppositions are not true, no conclusions can be drawn. Due to the many interactions between microbiome-targeted interventions and the microbiome, and the uniqueness of the microbiome, these presuppositions may not always be true. Moreover, as the previous point shows, it may be unrealistic to stratify for microbiome composition and, as argued in our previous article (Zeilstra *et al.*, 2018), a more personalized approach can help to overcome this problem.

Fourth, as it is known that multiple microbial genes or e.g. bacterial species can exert the same function,

assessing functional properties of the microbiome can be expected to be more useful in practice than attempting to link taxonomic or genomic composition to health problems (Heintz-Buschart and Wilmes, 2018). One reason is that genes must be expressed for the functional attributes to be expressed. In other words, genomically identical bacteria can have different phenotypes and thus functions. Moreover, as argued in the second point, a more pragmatic approach could be to address functions at higher level rather than attempting to understand functionality by e.g. mapping the functions of every single metabolite. High-level functional properties that have been identified include the butyrate-producing capacity, synthesis of essential vitamins, the removal of toxic compounds, the out-competition of pathogens, the strengthening of the intestinal barrier, and the stimulation and regulation of the immune system (Heintz-Buschart and Wilmes, 2018). These include well-known and classical mechanisms of action of probiotics (Cunningham *et al.*, 2021; Plaza-Diaz *et al.*, 2019; Sanders *et al.*, 2019).

Lastly, most pharmaceutical interventions are often specifically developed and studied for a single health problem, although it is not uncommon that pharmaceutical products are sometimes investigated for other health problems than their original scope as well. However, because the microbiome is associated with so many different health-issues, microbiome-targeted interventions, such as specific probiotic products, have often been studied for multiple, at first glance unrelated health problems. As examples, the strain *Lactiplantibacillus plantarum* DR7 have been shown in RCTs to reduce both symptoms of stress and anxiety, as well as upper respiratory tract infections (Altadill *et al.*, 2021; Chong *et al.*, 2019a), and the probiotic yeast *Saccharomyces boulardii* has shown in meta-analyses of RCTs to reducing the risk of antibiotic-associated diarrhoea as well as improve eradication rates of *Helicobacter pylori* infection when added to standard therapy (Szajewska *et al.*, 2015; Szajewska and Kołodziej, 2015). Specific strains may be involved in mechanisms of action that are sometime still unknown but possibly play a role in various diverse health problems. From a pragmatic point of view, this means that when an multimorbid patient suffers from two or more health problems for which studies of a specific microbiome-targeting intervention (e.g. a probiotic strain) have found evidence of an effect, it is reasonable to expect that this particular intervention can be beneficial in that individual patient.

In summary, these aspects show that microbiome-targeting interventions are excellent candidates for a pragmatic approach that integrates both EBM and PM and in the next section we describe the development and use in clinical practice of such an approach as a practical example.

8 A patient-centred approach

The aforementioned factors explain that translating evidence from studies to real-life situations in order to assess whether or not a particular patient will benefit from a certain intervention is, in many cases, very difficult. All three pillars of EBM are essential for clinical decision making: evidence from scientific research, individual clinical expertise, and taking into account patients' preferences. An important obstacle that makes integration of these pillars difficult is multimorbidity. While in the daily practice the first pillar of EBM (scientific clinical evidence) is input for the second pillar (employing clinical expertise), the concept of 'real-world evidence' enables the reverse route as well. Real-world evidence can come from different sources, for example pragmatic clinical trials, or real-world extension studies (Burcu *et al.*, 2022; Baumfeld Andre *et al.*, 2020), but in any case includes so called 'real-world data'. We understand real-world data here as defined by Klonoff (2019): 'information gathered through observations of routine clinical practice from multiple sources that can be linked together to provide meaningful patterns' and when this is systematically analysed it can yield real-world evidence. In our view, this includes patients as one of the multiple sources that generate these data. This adds to evidence from scientific clinical studies, therewith providing expansion of the evidence-base in the first pillar of EBM. In addition to real-world data, phase IV studies, that continue investigation of interventions after market introduction (Glasser *et al.*, 2007), may also be used to connect the second pillar of EBM back to the first, although the applicability of this route is limited because the vast majority of these studies focus on safety aspects rather than efficacy (Zhang *et al.*, 2016).

One way to improve the classic route (pillar 1 providing input to pillar 2) and to enhance the reverse (pillar 2 providing input to pillar 1), is to place the patient's individual situation and preferences as well as a life-course perspective in the centre of the clinical decision making and to integrate this with PM. As previously described, from effectiveness perspective, the ideal study would be to study one patient and administer a broad number of interventions to this individual (Figure 2B), instead of

the usual other way around (one intervention administered to a large number of patients, Figure 2A). While practical, financial, medical, and ethical obstacles limit the possibility to conduct this in everyday practice, it is possible to take into account all comorbidities simultaneously, including known mechanisms as well as the patient's medical background and complaints, to decide upon an integrated intervention, rather than treating health problems individually as is common practice. That is, if a patient has T2DM, MDD, has received a number of courses of antibiotics in the past years, and experiences fatigue and GI complaints, like our example patient, typical medical practice will likely focus on treating T2DM separately (e.g. with metformin) and possibly MDD (e.g. with a selective serotonin reuptake inhibitor), without much attention to the medical history and complaints. It is to be expected that taking all these relevant aspects into account improves the chances of success of the interventions, as the interventions will be selected using more information, yielding a more targeted option that is a better match to the individual patient (in other words: personalized treatment). The development of systematic critical appraisal methods such as GRADE (Schünemann *et al.*, 2013), and computerized information processing techniques have enabled an approach in which the patient profile can be aligned to available evidence of a range of interventions in order to select the best match. We call this Evidence-based Personalized Medicine (EBPM).

9 Example: EBPM for microbiome-targeting treatments

An example application of EBPM has been developed for microbiome-targeting treatments within the MyOwnResearch project in The Netherlands, by a collaboration of experts from the field of immunology, pharmacology, computational sciences and citizen science, and clinicians (Health Holland, 2018). This example approach, called the Personalized Microbiome-targeting Intervention Approach (PMIA), uses an online platform (developed as theoretical framework in the MyOwnResearch project, with one actual implementation accessible via microbiome-center.nl) via which HCPs integrate the complaints of the particular patient with the (medical) diagnosis that have been established in combination with insights from faecal analysis results, to come up with a personalized intervention proposal (Figure 3). Interventions are based on a combination of live micro-organisms (probiotics or live biotherapeutic products) and other microbiome-targeting compounds (e.g. prebiotics). The personalized intervention can consist of either the best matching off-the-shelf food supplements, or of a personalized preparation consisting of specific ingredients that is prepared for the individual patient. Importantly, aspects such as dose or duration are based on what is used in clinical studies, as also is recommended for probiotics (Jackson *et al.*, 2019). In remainder of the text both options are collectively referred to as intervention components. The key question within the PMIA approach is how to select the most appropriate microbiome-targeting intervention component(s) and this is where the integration of EBM and PM comes in.

First, all available scientific evidence for all available intervention components is assessed and graded according to an approach based on the GRADE method (Schünemann *et al.*, 2013). In the grading approach, a number of steps are followed for each intervention component:

1. First, all available evidence for an intervention component is gathered using PubMed and Google Scholar. In most cases, the search is conducted using the specific probiotic strain name or formulation name without restrictions in study type.

2. Next, a list of medical indications is extracted from this available scientific evidence. Indications can include health problems but also known (microbiome-related) mechanisms such as pathogen inhibition or anti-inflammatory effects. The mechanisms that are identified are in line with the aforementioned known high-level functional properties of the gut microbiome (butyrate-producing capacity, the out-competition of pathogens, the strengthening of the intestinal barrier, etc.).

3. Third, for each intervention component all available evidence is separately assessed per indication. This results in a score for the grade of evidence per indication per intervention component. The evidence can range from *in vitro* research to meta-analyses of RCTs, with a corresponding range of baseline scores (1 for *in vitro* studies till 4 for meta-analyses of RCTs, Table 1). Per study a baseline score is assigned. Score modification is applied if there are reasons to do so in accordance with a pre-defined list of score modifications (Table 2 and Table 3). When multiple studies provide evidence for the same indication, the overall score is based on the scores of the individual studies, taking into account whether or not their results are consistent.

4. These final scores are stored in a database (Figure 3A) which is queried by an online advice tool that HCPs

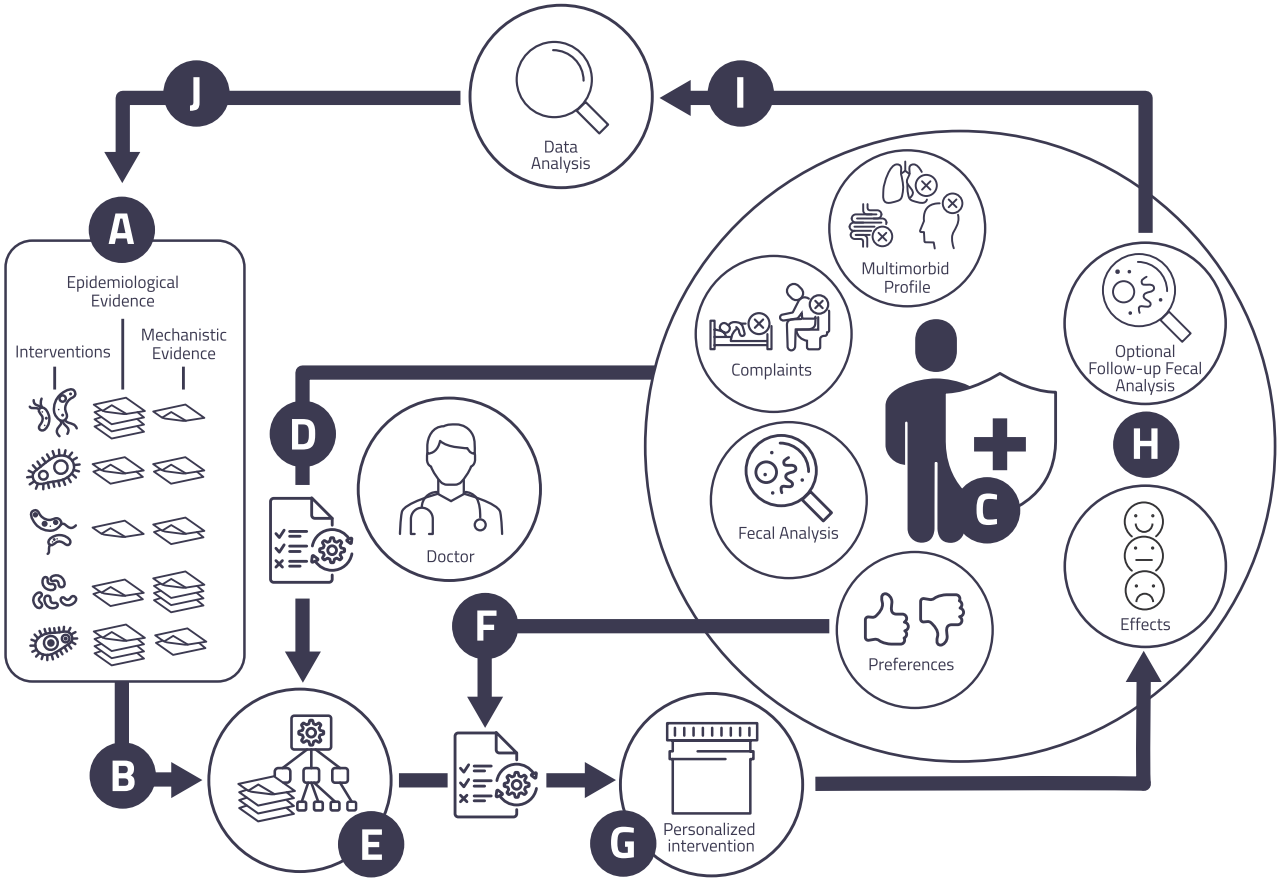

FIGURE 3 Overview of a pragmatic approach to integrate EBM and PM in the treatment of patients and gathering and understanding evidence. (A) For all intervention components (e.g. probiotic strains) all available epidemiological and mechanistic evidence is assessed and graded. The epidemiological evidence typically consists of clinical trials with efficacy of the intervention component on a particular health problem as outcome measure and are used by both EBM and PM. Mechanistic evidence can come from various study types and typically assesses effects on one or more mechanisms of action or pathways (e.g. immunological responses) and is typically used primarily in the PM approach. (B) For each individual intervention component, all evidence grades are entered into the algorithm (E). (C) All health aspects of each individual patient are taken into account and include the medical history and background (i.e. typically a multimorbid situation), the complaints, faecal analysis results, and the patient's preferences. For repeat treatments, the results (H) of previous treatment(s) may also be taken into account. (D) The patient profile is professionally assessed and entered into the system by the doctor using their clinical expertise. (E) An algorithm combines all evidence of all available intervention components with the patient profile to generate a suggested microbiome-targeting treatment. (F) The doctor assesses the suggested treatment using their clinical expertise and taken into account the patient's preferences and may adjust where they deem required. (G) The final selected personalized microbiome-targeting intervention is prescribed, prepared and administered. (H) The patient can provide feedback about the experienced effects. Optionally, a follow-up faecal analysis may be conducted. (I) Anonymized data from many patients, including their complaints, multimorbid situation, faecal analysis results, and effects are used in data analyses to continuously gain further understanding of potential subgroups and effects of intervention components. (J) The evidence from data analyses expands the evidence from scientific studies, is graded with the same strict methods, and is integrated into the approach via step A and B.

can use (see step 6). Thus, the database contains scores per intervention component per indication.

5. The fifth step of the approach is to periodically search for new evidence and update the scores accordingly.

The above listed steps provide a systematic means for assessment and grading of the first pillar of EBM, scientific evidence. Next, this is linked to the second and third pillar of EBM, clinical expertise and patient's preferences, towards a personalized intervention as follows:

6. An online tool has been developed in which doctors can enter characteristics of an individual patient for three categories: complaints, medical background and history, and mechanistic aspects that can be linked to faecal analysis results. The questions that are asked in the tool for each category are based on the current evidence for all available intervention components. That is,

TABLE 1 Baseline scores based on study types

Score	Evidence types	Rationale
1	· *In vitro* studies · Animal model studies	These two types of evidence provide the lowest level of certainty. Although animal studies do provide an *in vivo* test, their predictability for and translatability to humans is very poor (Kramer and Greek, 2018), and as such there are no convincing reasons to consider them to be of a higher reliability level than *in vitro* evidence.
1.5	· Case-reports · Collected observational data	Case reports provide evidence of *in vivo* use in humans but can be substantially influenced by individual circumstances. Collected observational data can be influenced by reversed causation, or by other interventions given simultaneously. Therefore, the baseline score is limited to the same level as case-reports.
2	· Case-series · Open label studies without control · Observational studies	Case series (if homogeneous) report similar outcomes in multiple case. Still, in case-series, studies without control, or observational studies there can be many confounding factors, and the setup is prone to bias, which cannot be controlled for.
2.5	· Case-control studies · Cohort studies	These study types include groups that were not exposed to the factor of interest, hence provide some form of control. However, they are of observational nature and do not control for confounding and bias.
3	· Non-blinded clinical trials	Intervention studies increase the reliability of the evidence due to the fact that the factor of interest is provided as an intervention. This eliminates some confounding. However, non-blinded trials do have a higher risk of bias than RCTs.
3.5	· Randomized controlled trials	These studies have the lowest risk of confounding and bias and therewith they offer the most reliable evidence that the intervention is causally related to the observed effect.
4	· Systematic reviews of RCTs · Meta-analyses of RCTs	If multiple RCTs with the same intervention show an unambiguous and repeatable effect, the reliability that this a true effect attributable to the intervention is very high.

if there is at least one intervention component for which there is evidence of an effect on constipation, but there is no evidence for any intervention component for sad mood, the complaints category will contain a question about constipation, but not a question about sad mood. If new intervention components are added, or new evidence becomes available, new questions can be added.

7. For each question, the HCP can provide a score that expresses the relevance and severity of each item (complaint, medical background, faecal analysis result (if available)) for the individual patient, therewith compiling a patient profile (Figure 3D). Importantly, the list of question and thus the patient profile takes into account many highly prevalent health problems and thus typically addresses a multimorbid situation. Moreover, the HCP together with the patient can adjust scores to express the patient's preferences about which health problems are more important for the patient to focus on. With regard to the third category, faecal analysis results, the questions are all based on known high-level functional aspects of the microbiome, such as for example butyrate-producing capacity, inhibition of pathogens, or intestinal barrier function. The faecal analysis types that HCPs can use through the platform are designed to enable easy interpretation of these high-level functional aspects by HCPs, and are not, as often seen with commercially available faecal analyses, merely a list of relative abundance of bacterial genera or species. The focus on high-level functional aspects ensures that the results of the faecal analyses are informative and useful for the clinical decision-making process, as disturbed functions can be used as targets for treatment.

8. Based on the patient profile and the graded evidence per intervention component, an algorithm com-

TABLE 2 List of factors that lower the baseline score

Factor	Effect	Rationale
A) Poor study design or conduct	–1 for levels 2 and higher	Poor study design has a larger impact on large scale observational and intervention studies. Case reports are hardly affected by a design.
B) Inconsistent results	–2 for levels 2.5 and higher –1 for levels 1 and higher	If results among different studies are inconsistent, this is a relatively strong indication that observed effects may not be casually related to the intervention.
C) Indirect evidence	–1 or –2 for levels 1.5 and higher, depending on the level of representativeness	Indirect evidence does not show that the effect is attributable to the intervention.
D) Imprecise outcomes	–1 for levels 1.5 and higher	If the outcomes are imprecise, it is less clear what the exact effect is and as such lowers the reliability that the effect is attributable to the intervention.
E) Publication bias	See: Inconsistent results	If there is publication bias, this may indicate that the actual outcomes are inconsistent.
F) Studied intervention differs from evaluated intervention	–1 or –2 for levels 1.5 and higher, depending on the level of representativeness For level 1: for large differences: –1.	This is a specific form of indirect evidence.
G) Low external validity	–1 for levels 1.5 and higher	If very narrowly defined eligibility criteria are used, the generalizability of the effect can be poor, which lowers the reliability that the effect (in general) is attributable to the intervention.
H) Very small sample	See: Low external validity	Small studies (e.g. n ≤ 30) are more prone to low generalizability.
I) Weak recommendation	–1	As per GRADE a weak recommendation means that the ratio between the likelihood of positive effects and the risk of adverse effects is low.
R) Small magnitude of clinical effect	–1 or –2 for levels 1.5 and higher, depending on the magnitude of clinical effect.	A small clinical effect magnitude is a reason for a weaker recommendation. Some rules of thumb (Livingston *et al.*, 2009): · For independent means of outcomes with continuous distribution SMD (Higgins and Green, 2011) can be used, with d ≤ 0.20 is considered small, d = 0.5 is medium and d ≥ 0.8 large. · For proportions Cohen's h can be used, with h ≤ 0.20 is considered small, h = 0.5 medium and h ≥ 0.8 large.

TABLE 3 List of factors that increase the baseline score

Factor	Effect	Rationale
J) Large magnitude of clinical effect	+1	A large clinical effect magnitude is a stronger indication of causality, as well as a reason for a stronger recommendation. See factor R for rules of thumb on effect sizes.
K) Dose-response gradient observed	+1	A dose-response effect is a stronger indication of causality.
L) Intervention is shown to directly and substantially affect known causal mechanism	+1	In some cases, disease mechanisms may have been shown to be causal. If evidence shows that an intervention has an effect on such causal mechanism, this increases the likelihood that the intervention will have a clinical effect. An example is the causal role of butyrate vs. the associative role of Bifidobacteria.
M) Different studies show consistent effect	+1 for levels 1.5 and higher	If different studies, with different measurement methods and different human participants consistently show the same effect, this is a stronger indication of causality.
N) Strong recommendation	+1	As per GRADE a strong recommendation means that the ratio between the likelihood of positive effects and the risk of adverse effects is high.

piles an advice for (a composition of) a personalized intervention for which the combined evidence shows that the intervention component(s) can address the multiple health problems of the patient (Figure 3E). The personalized intervention can consist of a personalized mixture of intervention components, or of a single intervention component. The algorithm will put intervention components higher on the list if there is evidence for a beneficial effect on multiple items from the patient profile (e.g. IBS, sad mood, and overgrowth of parasites), and if the patient scores are higher (based on severity and/or patient preferences). Thus, the proposal does not consist of one intervention component per indication (complaint, medical background, faecal analysis result), but typically consist of intervention components for which there is evidence for multiple indications. This way, intervention components are more likely to target common underlying mechanisms. This is even the case when the common underlying mechanisms for comorbid health problems (e.g. IBS and sad mood) are unknown, as intervention components may be selected based only on the fact that there is clinical evidence for distinct health problems, without knowledge about the mechanisms.

9. For each proposed intervention component, the HCP can see the matches between the grade-score and the patient profile and the bibliography of the publications for each indication.

With this approach the HCP gets an advice that is based on a critical appraisal of all available literature, on the multimorbid profile of the patient, as well as on the patient preferences. This approach partially automates combining the two elements of EBM: evidence from scientific studies and individual clinical expertise. The HCP assesses, based on their clinical expertise, the severity of all complaints, medical background items, and faecal analysis results of an individual patient, and the system presents the intervention components with evidence that best matches this patient profile and preferences. The system presents a concise description of the scientific evidence and a bibliography with all references, which drastically reduces the time that the physician needs to identify and assess relevant literature.

To make this more tangible, we use our hypothetical patient as an example. In step (7), the HCP together with this 60 year old female patient can score in the complaints section a relatively high score for fatigue, as well as for anxiety and for depression. Moreover, together they can decide how high to score for abdominal pain and bloating, and whether the varying stool consistency warrants scores for diarrhoea and/or constipation. In the second section, covering the medical background, the HCP can score urinary tract infections, metabolic dysfunction, and overweight. In addition, with the GI complaints in mind, the HCP could score either IBS and/or SIBO. This is an example where clinical expertise is crucial and well-experienced or well-trained HCPs can recognize that fatty stools (steatorrhea), and bloating and belching (especially after a meal), together with the other GI complaints, can be signs of SIBO (Quigley *et al.*, 2020). At the time of writing, no evidence was

available for any of the ingredients that HCPs can use for kidney stones or hypertension, so these indications cannot be filled out. Over time, new ingredients or new evidence may becomes available that do have evidence for these indications, and questions can be added to the list. For our example patient a faecal analysis was available at functional level, and in the third set of questions this can be filled in. For this example, a slightly low level of butyrate producing capacity is filled in, as well as some overgrowth of pathogenic bacteria, a slightly lowered level of mucus production-stimulating bacteria, a substantially elevated level of inflammation, some elevated markers of increased intestinal permeability, and substantial gluten intolerance.

With these scores, in step (8) a personalized intervention is drafted. Based on the scores and available evidence, this consists of partially hydrolysed guar gum (a prebiotic), *Lactiplantibacillus plantarum* DR7, *Bifidobacterium animalis subsp lactis* HN019, *Latilactobacillus sakei* probio65, a mixture of *Enterococcus faecium* Rosell-26 and *Bacillus subtilis* Rosell-179, and pasteurized *Akkermansia muciniphila* MucT. To take the *Lactiplantibacillus plantarum* DR7 as an example, this strain is proposed for our example patient based on her scores and the evidence it has for anxiety (e.g. Chong *et al.*, 2019a), inflammation (e.g. Chong *et al.*, 2019a,b), and metabolic dysfunction (primarily mechanistic evidence via AMPK modulation, see e.g. Desjardins and Steinberg, 2018; Lew *et al.*, 2018; Yap *et al.*, 2020). As mentioned before, the concept is not based on one ingredient for one indication, but rather uses a cross networked approach, in which several ingredients are proposed with overlapping effects (based on the available evidence). In this example, partially hydrolysed guar gum is proposed based on, among others, evidence for effect on metabolic dysfunction (i.e. improving insulin sensitivity, see e.g. Chuang *et al.*, 1992; Dall'Alba *et al.*, 2013; Peterson *et al.*, 1987), but also some evidence on anxiety and depressed mood (Parisi *et al.*, 2005), and evidence for its potential to improve abdominal pain (Romano *et al.*, 2013). For all proposed ingredients a similar partially overlapping set of indications applies, based on evidence ranging from stronger to weaker. The complete set of evidence on which the proposed personalized mixture is based, is visible for the HCP for each ingredient, as per step (9).

Since the systematic approach includes periodic updates of the assessment of published scientific evidence for each intervention component, another practical aspect is taken into account as well. That is, the rate in which new scientific evidence is gathered and published is ever increasing and by and large exceeds the rate at which new insights are implemented in clinical practice. Especially in the area targeting the gut microbiome this is not merely a luxury, but a necessity as in 2020 over 24,000 mechanistic and intervention studies were published. The approach described above enables continuous additions of new intervention components as well as new scientific evidence, aiding more rapid implementation in clinical practice of new findings, but with the safeguard of the systematic critical appraisal process. Moreover, it allows for the adoption of new microbiome analyses techniques, such as shotgun metagenomics, when these become cost-effective. Compared to amplicon sequencing, metagenome sequencing may reveal more information about low abundance taxa and can provide insights into metabolic activity of microbes (Durazzi *et al.*, 2021; Malla *et al.*, 2019). These aspects can help to get an enriched picture of the functional activity of the microbiome, hence further improve the targeting of treatments.

Real-world evidence from EBPM for microbiome-targeting treatments

In addition to aiding HCPs in treating their patients with a personalized microbiome-targeting intervention, the computerized information processing techniques and digital infrastructure of the aforementioned EBPM example allows for the reverse route as well: using real-world evidence to expand the evidence-base. Systematic consultation of the effects experienced by patients is conducted via phone calls. In addition, during and for some time after using a personalized microbiome-targeting intervention, patients are asked every few weeks to score on the same platform that HCPs can use the patient-reported complaints that are relevant to the particular individual (Figure 3H). Moreover, with patient consent anonymized data on the health problems (patient profile) and faecal analysis results can be used for data analyses in order to identify responders. Other means of gathering, analysing, and integrating this real-world data can be employed too, for example by analysing voluntarily filled out diaries (Van der Geest *et al.*, 2021). Together, these real-world data can be used to identify patterns and expand the evidence available for the intervention components (Figure 3I and 3J). This kind of real-world evidence not only enables the reverse route, from step 2 of EBM back to step 1, but, if it is used in the aforementioned procedure, it also creates a cycle. This cycle is much more rapid than the traditional route of clinical research, although it comes with caveats. For example, this route does not include blind-

ing or placebos and may be prone to e.g. reversed causation. Therefore, and because this kind of real-world evidence is not meaningless, the baseline score of this type of real-world data is equal to that of case reports (Table 1).

10 Discussion

A major challenge in the daily practice of clinicians is to determine which intervention is best for an individual patient. In the past decades, various paradigms have arisen that attempt to improve this using scientific research. The most prominent paradigms are EBM and PM. While EBM consists of three pillars (evidence from systematic research, clinical expertise, and patients' preferences), often most attention is drawn to evidence from systematic research, in particular RCTs. However, evidence from clinical studies is not always that informative for the clinical decision-making process for a particular patient. Although RCTs are valuable methods to investigate the impact of certain interventions and provide the least biased means to assess their efficacy, this type of research has some drawbacks as the study population often is not representative for the patient population seen in the practice due to narrow inclusion criteria used for studies and, in the vast majority of studies, not taking into account (or even exclusion of) the most prevalent chronic condition, multimorbidity. Moreover, the average effect found in clinical trials does often not predict the outcome of an intervention for an individual patient, just as the average shoe size will not fit for most people. The PM paradigm strives to overcome the latter by identifying responder-discriminating traits in clinical trials and use these to better predict the expected outcome for an individual patient.

Both paradigms are based on a realist worldview and presuppose that insights from structured research can be generalized to the real-world patient population, although both paradigms approach this in a very different way. The original EBM definition did embrace clinical expertise and, notably, mentions individual clinical expertise first (Sackett *et al.*, 1996). Moreover, the PM paradigm acknowledges that, with all new techniques to assess patient characteristics that might determine intervention effects, integration of various expertise has become even more important (Mason-Suares *et al.*, 2016). Therefore, we plea for a more pragmatic approach that acknowledges that interpretation of scientific evidence is by definition coloured by presuppositions and that no study can provide 100% certainty about the effect of an intervention in an individual patient. The approach that we propose attempts to overcome the epistemological arguments that seemingly divide the EBM and PM paradigms, and to use the strengths of both approaches, including patients' preferences, and acknowledges the limitations of both as well, in order to find a pragmatic way to improve clinical decision-making. That is, the proposed approach uses both insights of average effects from clinical trials, as well as insights about traits that may discriminate responders from mechanistic studies and integrates that into a model focused on multimorbidity. We call this pragmatic approach Evidence-based Personalized Medicine (EBPM).

With the gut microbiome being associated with almost all chronic diseases, the microbiome being highly individual, and many interventions being studied for a diverse range of different health problems, we have used microbiome-targeted interventions as a model for EBPM. The example EBPM for microbiome-targeting treatments integrates scientific evidence from different study types and grades it accordingly as weaker or stronger, offers a means to tailor an intervention to a multimorbid patient including their preferences, and uses known high-level functional properties of the gut microbiome as one means to further personalize the intervention. Moreover, it offers a way for HCPs to incorporate patients' preferences by emphasizing health problems that are more important for the patient. In addition, the proposed approach offers a means to feedback patients' experiences (real-world evidence) into the algorithm, to more rapidly learn what works for who, therewith closing the learning cycle. The PMIA example presented here is one implementation of EBPM, but other implementations, such as a flowchart-based approach, could be used too. We believe that EBPM can benefit patients with other intervention types as well, including nutritional interventions, and provides a means of fully utilizing all pillars of EBM as well as the strengths of PM.

Acknowledgements

The authors would like to thank and acknowledge all members of the MyOwnResearch project for the contribution to the ideas that are further developed in this article. The MyOwnResearch project (in full: *MyOwnResearch*: *Homogeneous subgroup identification in fatigue management across chronic autoimmune syn-*

dromes through Single Subject Study design) received funding from Health~Holland Topsector Life Sciences and Health and the Dutch Collaborating Health Funds (SGF).

Authors' contribution

Conceptualization, DZ; writing-original draft preparation, DZ; literature search: all; writing- review and editing, all; visualization, DZ; supervision, DZ, RB, AK; project administration, DZ. All authors have read and approved the final manuscript.

Conflict of interest

DZ is affiliated with Microbiome Center. IB was employee at Winclove Probiotics at the time of writing. RJB has received consulting fees from Akkermansia Company, Belgium, and research grants from AnaBio Technologies, Ireland and Biogaia AB, Sweden. None of the other authors have any interests to declare.

Funding

This work is a result of the MyOwnResearch project (in full: MyOwnResearch: Homogeneous subgroup identification in fatigue management across chronic autoimmune syndromes through Single Subject Study design). The MyOwnResearch project received funding from Health~Holland Topsector Life Sciences and Health and the Dutch Collaborating Health Funds (SGF)

References

Abdallah, M.S., Mosalam, E.M., Zidan, A.-A.A., Elattar, K.S., Zaki, S.A., Ramadan, A.N. and Ebeid, A.M., 2020. The antidiabetic metformin as an adjunct to antidepressants in patients with major depressive disorder: a proof-of-concept, randomized, double-blind, placebo-controlled trial. Neurotherapeutics 17: 1897-1906. https://doi.org/10.1007/s13311-020-00878-7

Abdallah, M.S., Mosalam, E.M., Zidan, A.-A.A., Elattar, K.S., Zaki, S.A., Ramadan, A.N. and Ebeid, A.M., 2022. Retraction note: the antidiabetic metformin as an adjunct to antidepressants in patients with major depressive disorder: a proof-of-concept, randomized, double-blind, placebo-controlled trial. Neurotherapeutics 19: 1687. https://doi.org/10.1007/s13311-022-01291-y

Altadill, T., Espadaler-Mazo, J. and Liong, M.-T., 2021. Effects of a lactobacilli probiotic on reducing duration of URTI and fever, and use of URTI-associated medicine: a re-analysis of a randomized, placebo-controlled study. Microorganisms 9: 528. https://doi.org/10.3390/microorganisms9030528

Ames Room (Philip Zimbardo), 2010. Available at: https://www.youtube.com/watch?v=hCV2Ba5wrcs.

Baumfeld Andre, E., Reynolds, R., Caubel, P., Azoulay, L. and Dreyer, N.A., 2020. Trial designs using real-world data: The changing landscape of the regulatory approval process. Pharmacoepidemiology and Drug Safety 29: 1201-1212. https://doi.org/10.1002/pds.4932

Berg, G., Rybakova, D., Fischer, D., Cernava, T., Vergès, M.-C.C., Charles, T., Chen, X., Cocolin, L., Eversole, K., Corral, G.H., Kazou, M., Kinkel, L., Lange, L., Lima, N., Loy, A., Macklin, J.A., Maguin, E., Mauchline, T., McClure, R., Mitter, B., Ryan, M., Sarand, I., Smidt, H., Schelkle, B., Roume, H., Kiran, G.S., Selvin, J., De Souza, R.S.C., Van Overbeek, L., Singh, B.K., Wagner, M., Walsh, A., Sessitsch, A. and Schloter, M., 2020. Microbiome definition re-visited: old concepts and new challenges. Microbiome 8: 103. https://doi.org/10.1186/s40168-020-00875-0

Blackstone, E.H., 2019. Precision medicine versus evidence-based medicine. Circulation 140: 1236-1238. https://doi.org/10.1161/CIRCULATIONAHA.119.043014

Blane, D., Burton-Jeangros, C., Cullati, S. and Sacker, A. (eds.), 2015. A life course perspective on health trajectories and transitions. Springer International Publishing, Cham, Switzerland. https://doi.org/10.1007/978-3-319-20484-0

Bourgonje, A.R., Von Martels, J.Z.H., De Vos, P., Faber, K.N. and Dijkstra, G., 2018. Increased fecal calprotectin levels in Crohn's disease correlate with elevated serum Th1- and Th17-associated cytokines. PLoS ONE 13: e0193202. https://doi.org/10.1371/journal.pone.0193202

Briggs Early, K. and Stanley, K., 2018. Position of the Academy of Nutrition and Dietetics: the role of medical nutrition therapy and registered dietitian nutritionists in the prevention and treatment of prediabetes and type 2 diabetes. Journal of the Academy of Nutrition and Dietetics 118: 343-353. https://doi.org/10.1016/j.jand.2017.11.021

Burcu, M., Manzano-Salgado, C.B., Butler, A.M. and Christian, J.B., 2022. A framework for extension studies using real-world data to examine long-term safety and effectiveness. Therapeutic Innovation & Regulatory Science 56: 15-22. https://doi.org/10.1007/s43441-021-00322-8

Burns, P.B., Rohrich, R.J. and Chung, K.C., 2011. The levels of evidence and their role in evidence-based medicine. Plastic and Reconstructive Surgery 128: 305-310. https://doi.org/10.1097/PRS.0b013e318219c171

Bush, C.L., Blumberg, J.B., El-Sohemy, A., Minich, D.M., Ordovás, J.M., Reed, D.G. and Behm, V.A.Y., 2020. Toward the definition of personalized nutrition: a proposal by the american nutrition association. Journal of the American College of Nutrition 39: 5-15. https://doi.org/10.1080/07315724.2019.1685332

Casellas, F., Guise, C.H., Robles, V., Navarro, E. and Borruel, N., 2017. Patient preferences for inflammatory bowel disease treatment objectives. Digestive and Liver Disease 49: 152-156. https://doi.org/10.1016/j.dld.2016.09.009

Chakravartty, A., 2017. Scientific Realism. In: Zalta, E.N. (ed.) The Stanford Encyclopedia of Philosophy. Metaphysics Research Lab, Stanford University. Available at: https://plato.stanford.edu/archives/sum2017/entries/scientific-realism/.

Chan, K.L., Cathomas, F. and Russo, S.J., 2019. Central and peripheral inflammation link metabolic syndrome and major depressive disorder. Physiology 34: 123-133. https://doi.org/10.1152/physiol.00047.2018

Chong, H.-X., Yusoff, N.A.A., Hor, Y.-Y., Lew, L.-C., Jaafar, M.H., Choi, S.-B., Yusoff, M.S.B., Wahid, N., Abdullah, M.F.I.L., Zakaria, N., Ong, K.-L., Park, Y.-H. and Liong, M.-T., 2019a. *Lactobacillus plantarum* DR7 alleviates stress and anxiety in adults: a randomised, double-blind, placebo-controlled study. Beneficial Microbes 10: 355-373. https://doi.org/10.3920/BM2018.0135

Chong, H.-X., Yusoff, N.A.A., Hor, Y.-Y., Lew, L.-C., Jaafar, M.H., Choi, S.-B., Yusoff, M.S.B., Wahid, N., Abdullah, M.F.I.L., Zakaria, N., Ong, K.-L., Park, Y.-H. and Liong, M.-T., 2019b. *Lactobacillus plantarum* DR7 improved upper respiratory tract infections via enhancing immune and inflammatory parameters: A randomized, double-blind, placebo-controlled study. Journal of Dairy Science 102: 4783-4797. https://doi.org/10.3168/jds.2018-16103

Christovich, A. and Luo, X.M., 2022. Gut microbiota, leaky gut, and autoimmune diseases. Frontiers in Immunology 13: 946248. https://doi.org/10.3389/fimmu.2022.946248

Chuang, L.M., Jou, T.S., Yang, W.S., Wu, H.P., Huang, S.H., Tai, T.Y. and Lin, B.J., 1992. Therapeutic effect of guar gum in patients with non-insulin-dependent diabetes mellitus. Journal of the Formosan Medical Association = Taiwan Yi Zhi 91: 15-19.

Council conclusions on personalised medicine for patients, 2015. Available at: https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=uriserv%3AOJ.C_.2015.421.01.0002.01.ENG&toc=OJ%3AC%3A2015%3A421%3AFULL.

Cunningham, M., Azcarate-Peril, M.A., Barnard, A., Benoit, V., Grimaldi, R., Guyonnet, D., Holscher, H.D., Hunter, K., Manurung, S., Obis, D., Petrova, M.I., Steinert, R.E., Swanson, K.S., van Sinderen, D., Vulevic, J. and Gibson, G.R., 2021. Shaping the future of probiotics and prebiotics. Trends in Microbiology 29: 667-685. https://doi.org/10.1016/j.tim.2021.01.003

Dall'Alba, V., Silva, F.M., Antonio, J.P., Steemburgo, T., Royer, C.P., Almeida, J.C., Gross, J.L. and Azevedo, M.J., 2013. Improvement of the metabolic syndrome profile by soluble fibre - guar gum - in patients with type 2 diabetes: a randomised clinical trial. British Journal of Nutrition 110: 1601-1610. https://doi.org/10.1017/S0007114513001025

Dekkers, O.M., von Elm, E., Algra, A., Romijn, J.A. and Vandenbroucke, J.P., 2010. How to assess the external validity of therapeutic trials: a conceptual approach. International Journal of Epidemiology 39: 89-94. https://doi.org/10.1093/ije/dyp174

Desjardins, E.M. and Steinberg, G.R., 2018. Emerging role of AMPK in brown and beige adipose tissue (BAT): implications for obesity, insulin resistance, and type 2 diabetes. Current Diabetes Reports 18: 80. https://doi.org/10.1007/s11892-018-1049-6

Durazzi, F., Sala, C., Castellani, G., Manfreda, G., Remondini, D. and De Cesare, A., 2021. Comparison between 16S rRNA and shotgun sequencing data for the taxonomic characterization of the gut microbiota. Scientific Reports 11: 3030. https://doi.org/10.1038/s41598-021-82726-y

Eichelberger, M., Hagen, H., Kedrowski, M., Kransz, N., Shah, S. and Taylor, N., 2021. Patients want to feel better. who should pay for that? BCG Global Available at: https://www.bcg.com/publications/2021/evaluating-coverage-quality-of-life-benefits.

Emmert-Streib, F., 2013. Personalized medicine: Has it started yet? A reconstruction of the early history. Frontiers in Genetics 3: 313. https://doi.org/10.3389/fgene.2012.00313

Ernst, E. and Pittler, M.H., 2006. Efficacy or effectiveness? Journal of Internal Medicine 260: 488-490. https://doi.org/10.1111/j.1365-2796.2006.01707.x

Fassarella, M., Blaak, E.E., Penders, J., Nauta, A., Smidt, H. and Zoetendal, E.G., 2020. Gut microbiome stability and resilience: elucidating the response to perturbations in order to modulate gut health. Gut 70: 595-605. https://doi.org/10.1136/gutjnl-2020-321747

Felder, M., van de Bovenkamp, H., Meerding, W.J. and de Bont, A., 2021. Who contextualises clinical epidemiological evidence? A political analysis of the problem of evidence-based medicine in the layered Dutch healthcare system. Health Policy 125: 34-40. https://doi.org/10.1016/j.healthpol.2020.09.006

Franzosa, E.A., Huang, K., Meadow, J.F., Gevers, D., Lemon, K.P., Bohannan, B.J.M. and Huttenhower, C., 2015. Identifying personal microbiomes using metagenomic codes. Proceedings of the National Academy of Sciences of the USA 112: E2930-E2938. https://doi.org/10.1073/pnas.1423854112

Galea, S., Riddle, M. and Kaplan, G.A., 2010. Causal thinking and complex system approaches in epidemiology. International Journal of Epidemiology 39: 97-106. https://doi.org/10.1093/ije/dyp296

Gauch, H.G., 2003. Scientific method in practice. Cambridge University Press, New York, NY, USA, 435 pp.

Glasser, S.P., Salas, M. and Delzell, E., 2007. Importance and challenges of studying marketed drugs: what is a phase IV study? Common clinical research designs, registries, and self-reporting systems. Journal of Clinical Pharmacology 47: 1074-1086. https://doi.org/10.1177/0091270007304776

Gomaa, E.Z., 2020. Human gut microbiota/microbiome in health and diseases: a review. Antonie van Leeuwenhoek 113: 2019-2040. https://doi.org/10.1007/s10482-020-01474-7

Green, L.W., 2008. Making research relevant: if it is an evidence-based practice, where's the practice-based evidence? Family Practice 25: i20-i24. https://doi.org/10.1093/fampra/cmn055.

Greenhalgh, T., 2020. Will COVID-19 be evidence-based medicine's nemesis? PLOS Medicine 17: e1003266. https://doi.org/10.1371/journal.pmed.1003266

Guerrini, C.J., Kempner, J., Rasmussen, L.M. and Wexler, A., 2022. Diverse, emergent, disruptive: perspectives on and developments in biomedical citizen science. Citizen Science – Theory and Practice 7: 36. https://doi.org/10.5334/cstp.598

Guyatt, G., 1992. Evidence-Based medicine: a new approach to teaching the practice of medicine. JAMA 268: 2420. https://doi.org/10.1001/jama.1992.03490170092032

Guyatt, G.H., Haynes, R.B., Jaeschke, R.Z., Cook, D.J., Green, L., Naylor, C.D., Wilson, M.C., Richardson, W.S. and Group for the E.-B.M.W., 2000. Users' guides to the medical literature: XXV. Evidence-based medicine: principles for applying the users' guides to patient care. JAMA 284: 1290-1296. https://doi.org/10.1001/jama.284.10.1290

Hamer, J.A., Testani, D., Mansur, R.B., Lee, Y., Subramaniapillai, M. and McIntyre, R.S., 2019. Brain insulin resistance: A treatment target for cognitive impairment and anhedonia in depression. Experimental Neurology 315: 1-8. https://doi.org/10.1016/j.expneurol.2019.01.016

Haynes, R.B., Devereaux, P.J. and Guyatt, G.H., 2002. Clinical expertise in the era of evidence-based medicine and patient choice. BMJ Evidence-Based Medicine 7: 36-38. https://doi.org/10.1136/ebm.7.2.36

Health Holland, 2018. Top Sector LSH and Health Funds support the project MyOwnResearch with 2,6 million. Health Holland. Available at: https://www.health-holland.com/news/2018/06/myownresearch-beter-gezond-project.

Heintz-Buschart, A. and Wilmes, P., 2018. Human gut microbiome: function matters. Trends in Microbiology 26: 563-574. https://doi.org/10.1016/j.tim.2017.11.002

Heinzel, S., Jureczek, J., Kainulainen, V., Nieminen, A.I., Suenkel, U., von Thaler, A.-K., Kaleta, C., Eschweiler, G.W., Brockmann, K., Aho, V.T.E., Auvinen, P., Maetzler, W., Berg, D. and Scheperjans, F., 2024. Elevated fecal calprotectin is associated with gut microbial dysbiosis, altered serum markers and clinical outcomes in older individuals. Scientific Reports 14: 13513. https://doi.org/10.1038/s41598-024-63893-0

Heneghan, C. and Mahtani, K.R., 2019. Absolute effects of statins in the elderly. BMJ Evidence-Based Medicine 24: 200-202. https://doi.org/10.1136/bmjebm-2019-111189

Hernán, M.A. and Robins, J.M., 2022. Causal inference: what if. Chapman & Hall/CRC, Boca Raton, FL, USA. Available at: https://www.hsph.harvard.edu/miguel-hernan/causal-inference-book/.

Hernández Medina, R., Kutuzova, S., Nielsen, K.N., Johansen, J., Hansen, L.H., Nielsen, M. and Rasmussen, S., 2022. Machine learning and deep learning applications in microbiome research. ISME Communications 2: 98. https://doi.org/10.1038/s43705-022-00182-9

Higgins, J. and Green, S., 2011. Cochrane Handbook for Systematic Reviews of Interventions. The Cochrane Collaboration. Available at: www.handbook.cochrane.org.

Hillegass, E., Lukaszewicz, K. and Puthoff, M., 2022. Role of physical therapists in the management of individuals at risk for or diagnosed with venous thromboembolism: evidence-based clinical practice guideline 2022. Physical Therapy 102: pzac057. https://doi.org/10.1093/ptj/pzac057

Hopewell, S., Loudon, K., Clarke, M.J., Oxman, A.D. and Dickersin, K., 2009. Publication bias in clinical trials due to statistical significance or direction of trial results. Cochrane Database of Systematic Reviews 2009: MR000006. https://doi.org/10.1002/14651858.MR000006.pub3

Hsueh, P.-Y.S., Dey, S., Das, S. and Wetter, T., 2017. Making sense of patient-generated health data for interpretable patient-centered care: the transition from 'more' to 'better'. In: Gundlapalli, A.V., Jaulent, M.-C. and Zhao, D. (eds.) MEDINFO 2017: Precision Healthcare through Informatics. IOS Press, Amsterdam, the Netherlands, pp. 113-117. https://doi.org/10.3233/978-1-61499-830-3-113

Illingworth, V. (ed.), 1991. The Penguin dictionary of physics. Penguin Books, London, UK, 544 pp.

Jackson, S.A., Schoeni, J.L., Vegge, C., Pane, M., Stahl, B., Bradley, M., Goldman, V.S., Burguière, P., Atwater, J.B. and Sanders, M.E., 2019. Improving end-user trust in the quality of commercial probiotic products. Frontiers in Microbiology 10: 739. https://doi.org/10.3389/fmicb.2019.00739

Julious, S.A. and Mullee, M.A., 1994. Confounding and Simpson's paradox. BMJ 309: 1480-1481. https://doi.org/10.1136/bmj.309.6967.1480

Jüni, P., Altman, D.G. and Egger, M., 2001. Assessing the quality of controlled clinical trials. BMJ 323: 42-46. https://doi.org/10.1136/bmj.323.7303.42

Kelley, J.M. and Kaptchuk, T.J., 2010. Group analysis versus individual response: The inferential limits of randomized controlled trials. Contemporary Clinical Trials 31: 423-428. https://doi.org/10.1016/j.cct.2010.07.003

Kelly, M.P., Heath, I., Howick, J. and Greenhalgh, T., 2015. The importance of values in evidence-based medicine. BMC Medical Ethics 16: 69. https://doi.org/10.1186/s12910-015-0063-3

Kennedy-Martin, T., Curtis, S., Faries, D., Robinson, S. and Johnston, J., 2015. A literature review on the representativeness of randomized controlled trial samples and implications for the external validity of trial results. Trials 16: 495. https://doi.org/10.1186/s13063-015-1023-4

Kent, D.M., Steyerberg, E. and van Klaveren, D., 2018. Personalized evidence based medicine: predictive approaches to heterogeneous treatment effects. BMJ 363: k4245. https://doi.org/10.1136/bmj.k4245

Klonoff, D.C., 2019. The expanding role of real-world evidence trials in health care decision making. Journal of Diabetes Science and Technology 14: 174-179. https://doi.org/10.1177/1932296819832653

Kramer, L.A. and Greek, R., 2018. Human stakeholders and the use of animals in drug development. Business and Society Review 123: 3-58. https://doi.org/10.1111/basr.12134

Kravitz, R.L., Duan, N.H. and Braslow, J., 2004. Evidence-based medicine, heterogeneity of treatment effects, and the trouble with averages. Milbank Quarterly 82: 661-687.

Legg, C. and Hookway, C., 2021. Pragmatism. In: Zalta, E.N. (ed.) The Stanford Encyclopedia of Philosophy. Metaphysics Research Lab, Stanford University, CA, USA. Available at: https://plato.stanford.edu/archives/sum2021/entries/pragmatism/.

Lew, L.-C., Choi, S.-B., Khoo, B.-Y., Sreenivasan, S., Ong, K.-L. and Liong, M.-T., 2018. *Lactobacillus plantarum* DR7 reduces cholesterol via phosphorylation of AMPK that down-regulated the mRNA expression of HMG-CoA reductase. Korean Journal for Food Science of Animal Resources 38: 350-361. https://doi.org/10.5851/kosfa.2018.38.2.350

Lin, B. and Wu, S., 2022. Digital transformation in personalized medicine with artificial intelligence and the internet of medical things. OMICS 26: 77-81. https://doi.org/10.1089/omi.2021.0037

Livingston, E.H., Elliot, A., Hynan, L. and Cao, J., 2009. Effect size estimation: a necessary component of statistical analysis. Archives of Surgery 144: 706. https://doi.org/10.1001/archsurg.2009.150

Lundh, A., Lexchin, J., Mintzes, B., Schroll, J.B. and Bero, L., 2017. Industry sponsorship and research outcome. Cochrane Database of Systematic Reviews 2: MR000033. https://doi.org/10.1002/14651858.MR000033.pub3

Lyra e Silva, N. de M., Lam, M.P., Soares, C.N., Munoz, D.P., Milev, R. and De Felice, F.G., 2019. Insulin resistance as a shared pathogenic mechanism between depression and type 2 diabetes. Frontiers in Psychiatry 10: 57. https://doi.org/10.3389/fpsyt.2019.00057

Mackey, A. and Bassendowski, S., 2017. The history of evidence-based practice in nursing education and practice. Journal of Professional Nursing 33: 51-55. https://doi.org/10.1016/j.profnurs.2016.05.009

Malla, M.A., Dubey, A., Kumar, A., Yadav, S., Hashem, A. and Abd Allah, E.F., 2019. Exploring the human microbiome: the potential future role of next-generation sequencing in disease diagnosis and treatment. Frontiers in Immunology 9: 2868. https://doi.org/10.3389/fimmu.2018.02868

Marley, J., 2000. Efficacy, effectiveness, efficiency. Australian Prescriber 23: 114-115. https://doi.org/10.18773/austprescr.2000.131

Mason-Suares, H., Sweetser, D.A., Lindeman, N.I. and Morton, C.C., 2016. Training the future leaders in personalized medicine. Journal of Personalized Medicine 6: 1. https://doi.org/10.3390/jpm6010001

McCoy, L.G., 2020. What makes medical knowledge: on the epistemological tensions between 'evidence-based medicine' and 'personalized medicine'. University of Toronto Medical Journal 97: 27-31.

Miyauchi, E., Shimokawa, C., Steimle, A., Desai, M.S. and Ohno, H., 2023. The impact of the gut microbiome on extra-intestinal autoimmune diseases. Nature Reviews Immunology 23: 9-23. https://doi.org/10.1038/s41577-022-00727-y

Montori, V.M., Brito, J.P. and Murad, M.H., 2013. The optimal practice of evidence-based medicine: incorporating patient preferences in practice guidelines. JAMA 310: 2503-2504. https://doi.org/10.1001/jama.2013.281422

Niiniluoto, I., 1991. Realism, relativism, and constructivism. Synthese 89: 135-162.

Nurk, S., Koren, S., Rhie, A., Rautiainen, M., Bzikadze, A.V., Mikheenko, A., Vollger, M.R., Altemose, N., Uralsky, L., Gershman, A., Aganezov, S., Hoyt, S.J., Diekhans, M., Logsdon, G.A., Alonge, M., Antonarakis, S.E., Borchers, M., Bouffard, G.G., Brooks, S.Y., Caldas, G.V., Chen, N.-C., Cheng, H., Chin, C.-S., Chow, W., de Lima, L.G., Dishuck, P.C., Durbin, R., Dvorkina, T., Fiddes, I.T., Formenti, G., Fulton, R.S., Fungtammasan, A., Garrison, E., Grady, P.G.S., Graves-Lindsay, T.A., Hall, I.M., Hansen, N.F., Hartley, G.A., Haukness, M., Howe, K., Hunkapiller, M.W., Jain, C., Jain, M., Jarvis, E.D., Kerpedjiev, P., Kirsche, M., Kolmogorov, M., Korlach, J., Kremitzki, M., Li, H., Maduro, V.V., Marschall, T., McCartney, A.M., McDaniel, J., Miller, D.E., Mullikin, J.C., Myers,

E.W., Olson, N.D., Paten, B., Peluso, P., Pevzner, P.A., Porubsky, D., Potapova, T., Rogaev, E.I., Rosenfeld, J.A., Salzberg, S.L., Schneider, V.A., Sedlazeck, F.J., Shafin, K., Shew, C.J., Shumate, A., Sims, Y., Smit, A.F.A., Soto, D.C., Sović, I., Storer, J.M., Streets, A., Sullivan, B.A., Thibaud-Nissen, F., Torrance, J., Wagner, J., Walenz, B.P., Wenger, A., Wood, J.M.D., Xiao, C., Yan, S.M., Young, A.C., Zarate, S., Surti, U., McCoy, R.C., Dennis, M.Y., Alexandrov, I.A., Gerton, J.L., O'Neill, R.J., Timp, W., Zook, J.M., Schatz, M.C., Eichler, E.E., Miga, K.H. and Phillippy, A.M., 2022. The complete sequence of a human genome. Science 376: 44-53. https://doi.org/10.1126/science.abj6987

Ogilvie, D., Adams, J., Bauman, A., Gregg, E.W., Panter, J., Siegel, K.R., Wareham, N.J. and White, M., 2020. Using natural experimental studies to guide public health action: turning the evidence-based medicine paradigm on its head. Journal of Epidemiology and Community Health 74: 203-208. https://doi.org/10.1136/jech-2019-213085

Ordovas, J.M., Ferguson, L.R., Tai, E.S. and Mathers, J.C., 2018. Personalised nutrition and health. BMJ 361: bmj.k2173. https://doi.org/10.1136/bmj.k2173

Pan, A., Lin, X., Hemler, E. and Hu, F.B., 2018. Diet and cardiovascular disease: advances and challenges in population-based studies. Cell Metabolism 27: 489-496. https://doi.org/10.1016/j.cmet.2018.02.017

Parisi, G., Bottona, E., Carrara, M., Cardin, F., Faedo, A., Goldin, D., Marino, M., Pantalena, M., Tafner, G., Verdianelli, G., Zilli, M. and Leandro, G., 2005. Treatment effects of partially hydrolyzed guar gum on symptoms and quality of life of patients with irritable bowel syndrome. a multicenter randomized open trial. Digestive Diseases and Sciences 50: 1107-1112. https://doi.org/10.1007/s10620-005-2713-7

Park, J.-E., Jeong, G.-H., Lee, I.-K., Yoon, Y.-R., Liu, K.-H., Gu, N. and Shin, K.-H., 2018. A pharmacometabolomic approach to predict response to metformin in early-phase type 2 diabetes mellitus patients. Molecules 23: 1579. https://doi.org/10.3390/molecules23071579

Peterson, D.B., Ellis, P.R., Baylis, J.M., Fielden, P., Ajodhia, J., Leeds, A.R. and Jepson, E.M., 1987. Low dose guar in a novel food product: improved metabolic control in non-insulin-dependent diabetes. Diabetic Medicine 4: 111-115. https://doi.org/10.1111/j.1464-5491.1987.tb00843.x

Plaza-Diaz, J., Ruiz-Ojeda, F.J., Gil-Campos, M. and Gil, A., 2019. Mechanisms of action of probiotics. Advances in Nutrition 10: S49-S66. https://doi.org/10.1093/advances/nmy063

Quigley, E.M.M., Murray, J.A. and Pimentel, M., 2020. AGA clinical practice update on small intestinal bacterial overgrowth: expert review. Gastroenterology 159: 1526-1532. https://doi.org/10.1053/j.gastro.2020.06.090

Revel-Muroz, A., Akulinin, M., Shilova, P., Tyakht, A. and Klimenko, N., 2023. Stability of human gut microbiome: Comparison of ecological modelling and observational approaches. Computational and Structural Biotechnology Journal 21: 4456-4468. https://doi.org/10.1016/j.csbj.2023.08.030

Romano, C., Comito, D., Famiani, A., Calamarà, S. and Loddo, I., 2013. Partially hydrolyzed guar gum in pediatric functional abdominal pain. World Journal of Gastroenterology 19: 235-240. https://doi.org/10.3748/wjg.v19.i2.235

Sackett, D.L., Rosenberg, W.M., Gray, J.A., Haynes, R.B. and Richardson, W.S., 1996. Evidence based medicine: what it is and what it isn't. BMJ 312: 71-72. https://doi.org/10.1136/bmj.312.7023.71

Sanders, M.E., Merenstein, D.J., Reid, G., Gibson, G.R. and Rastall, R.A., 2019. Probiotics and prebiotics in intestinal health and disease: from biology to the clinic. Nature Reviews Gastroenterology and Hepatology 16: 605-616. https://doi.org/10.1038/s41575-019-0173-3

Schünemann, H.J., Brożek, J.L., Guyatt, G. and Oxman, A.D., 2013. GRADE Handbook. Handbook for grading the quality of evidence and the strength of recommendations using the GRADE approach. Available at: https://gdt.gradepro.org/app/handbook/handbook.html.

Senn, S., 2018. Statistical pitfalls of personalized medicine. Nature 563: 619-621. https://doi.org/10.1038/d41586-018-07535-2

Siontis, K.C., Evangelou, E. and Ioannidis, J.P., 2011. Magnitude of effects in clinical trials published in high-impact general medical journals. International Journal of Epidemiology 40: 1280-1291. https://doi.org/10.1093/ije/dyr095

Sturmberg, J.P., Bennett, J.M., Martin, C.M. and Picard, M., 2017. 'Multimorbidity' as the manifestation of network disturbances. Journal of Evaluation in Clinical Practice 23: 199-208. https://doi.org/10.1111/jep.12587

Sturmberg, J.P. and Marcum, J.A., 2023. From cause and effect to causes and effects. Journal of Evaluation in Clinical Practice 30: 296-308. https://doi.org/10.1111/jep.13814

Subbiah, V., 2023. The next generation of evidence-based medicine. Nature Medicine 29: 49-58. https://doi.org/10.1038/s41591-022-02160-z

Szajewska, H., Horvath, A. and Kołodziej, M., 2015. Systematic review with meta-analysis: Saccharomyces boulardii supplementation and eradication of *Helicobacter pylori* infection. Alimentary Pharmacology & Therapeutics 41: 1237-1245. https://doi.org/10.1111/apt.13214

Szajewska, H. and Kołodziej, M., 2015. Systematic review with meta-analysis: *Saccharomyces boulardii* in the prevention of antibiotic-associated diarrhoea. Alimentary Pharmacology & Therapeutics 42: 793-801. https://doi.org/10.1111/apt.13344

Tait, G., 2005. The ADHD debate and the philosophy of truth. International Journal of Inclusive Education 9: 17-38. https://doi.org/10.1080/1360311042000299775

Tan, Y.Y., Papez, V., Chang, W.H., Mueller, S.H., Denaxas, S. and Lai, A.G., 2022. Comparing clinical trial population representativeness to real-world populations: an external validity analysis encompassing 43 895 trials and 5 685 738 individuals across 989 unique drugs and 286 conditions in England. The Lancet Healthy Longevity 3: e674-e689. https://doi.org/10.1016/S2666-7568(22)00186-6

te Meerman, S., Batstra, L., Freedman, J.E., Hoekstra, R. and Grietens, H., 2020. ADHD and brain anatomy: what do academic textbooks used in the Netherlands tell students? Children & Society 34: 136-150. https://doi.org/10.1111/chso.12362

Thurber, S., Sheehan, W. and Roberts, R.J., 2009. Attention deficit hyperactivity disorder and scientific epistemology. Dialogues in Philosophy, Mental and Neuro Sciences 2: 33-39.

Tierney, B.T., Yang, Z., Luber, J.M., Beaudin, M., Wibowo, M.C., Baek, C., Mehlenbacher, E., Patel, C.J. and Kostic, A.D., 2019. The landscape of genetic content in the gut and oral human microbiome. Cell Host & Microbe 26: 283-295.e8. https://doi.org/10.1016/j.chom.2019.07.008

Tinetti, M.E., Fried, T.R. and Boyd, C.M., 2012. Designing health care for the most common chronic condition – multimorbidity. JAMA 307: 2493-2494. https://doi.org/10.1001/jama.2012.5265

Tonelli, M.R. and Shirts, B.H., 2017. Knowledge for precision medicine: mechanistic reasoning and methodological pluralism. JAMA 318: 1649-1650. https://doi.org/10.1001/jama.2017.11914

Torres, N. and Tovar, A.R., 2021. The present and future of personalized nutrition. Revista de Investigación Clínica 73: 321-325. https://doi.org/10.24875/RIC.21000346

Toulmin, S., 1990. Cosmopolis: the hidden agenda of modernity. Free Press, New York, NY, USA. 228 pp.

Van Baarlen, P., Troost, F.J., van Hemert, S., van der Meer, C., de Vos, W.M., de Groot, P.J., Hooiveld, G.J.E.J., Brummer, R.-J.M. and Kleerebezem, M., 2009. Differential NF-kappaB pathways induction by *Lactobacillus plantarum* in the duodenum of healthy humans correlating with immune tolerance. Proceedings of the National Academy of Sciences of the USA 106: 2371-2376. https://doi.org/10.1073/pnas.0809919106

Van der Geest, A.M., Besseling-van der Vaart, I., Schellinger-de Goede, E.M., van der Waal, M.B., Claassen, E., Flach, J. and van de Burgwal, L.H.M., 2021. Multispecies probiotics promote perceived human health and wellbeing: insights into the value of retrospective studies on user experiences. Beneficial Microbes 12: 413-430. https://doi.org/10.3920/BM2020.0162

Vandeputte, D., De Commer, L., Tito, R.Y., Kathagen, G., Sabino, J., Vermeire, S., Faust, K. and Raes, J., 2021. Temporal variability in quantitative human gut microbiome profiles and implications for clinical research. Nature Communications 12: 6740. https://doi.org/10.1038/s41467-021-27098-7

World Health Organization (WHO), 2022. Non communicable diseases. Available at: https://www.who.int/news-room/fact-sheets/detail/noncommunicable-diseases.

Xiong, R., Aiken, E., Caldwell, R., Vernon, S.D., Kozhaya, L., Gunter, C., Bateman, L., Unutmaz, D. and Oh, J., 2025. AI-driven multi-omics modeling of myalgic encephalomyelitis/chronic fatigue syndrome. Nature Medicine 31: 2991-3001. https://doi.org/10.1038/s41591-025-03788-3

Yap, P.-G., Choi, S.-B. and Liong, M.-T., 2020. Allantoin, a potential metabolite that promotes AMPK phosphorylation and suppresses cholesterol biosynthesis via the mevalonate pathway and Bloch pathway. Applied Biochemistry and Biotechnology 191: 226-244. https://doi.org/10.1007/s12010-020-03265-2

Zeilstra, D., Younes, J.A., Brummer, R.J. and Kleerebezem, M., 2018. Perspective: fundamental limitations of the randomized controlled trial method in nutritional research: the example of probiotics. Advances in Nutrition 9: 561-571. https://doi.org/10.1093/advances/nmy046

Zhang, J. and Knight, R., 2023. Genomic mutations within the host microbiome: adaptive evolution or purifying selection. Engineering 20: 96-102. https://doi.org/10.1016/j.eng.2021.11.018

Zhang, X., Zhang, Y., Ye, X., Guo, X., Zhang, T. and He, J., 2016. Overview of phase IV clinical trials for postmarket drug safety surveillance: a status report from the ClinicalTrials.gov registry. BMJ Open 6: e010643. https://doi.org/10.1136/bmjopen-2015-010643