Science & Education (2023) 32:1709–1738
https://doi.org/10.1007/s11191-022-00363-x

ARTICLE

Integration of Philosophy of Science in Biomedical Data Science Education to Foster Better Scientific Practice

Annelies Pieterman-Bos[1] · **Marc H. W. van Mil**[1,2]

Accepted: 3 July 2022 / Published online: 23 July 2022

Abstract

Biomedical data science education faces the challenge of preparing students for conducting rigorous research with increasingly complex and large datasets. At the same time, philosophers of science face the challenge of making their expertise accessible for scientists in such a way that it can improve everyday research practice. Here, we investigate the possibility of approaching these challenges together. In current and proposed approaches to biomedical data science education, we identify a dominant focus on only one aspect of conducting scientific research: understanding and using data, research methods, and statistical methods. We argue that this approach cannot solve biomedical data science's challenge and we propose to shift the focus to four other aspects of conducting research: making and justifying decisions in research design and implementation, explaining their epistemic and non-epistemic effects, balancing varying responsibilities, and reporting scientific research. Attending to these aspects requires learning on different dimensions than solely learning to apply techniques (first dimension). It also requires learning to make choices (second dimension) and to understand the rationale behind choices (third dimension). This could be fostered by integrating philosophical training in biomedical data science education. Furthermore, philosophical training fosters a fourth dimension of learning, namely, understanding the nature of science. In this article, we explain how we identified the five aspects of conducting research and the four dimensions of learning, and why attending to the fourth dimension is essential. We discuss educational approaches to attend to all aspects and dimensions, and present initial design principles to implement these approaches.

✉ Annelies Pieterman-Bos
j.e.bos-11@umcutrecht.nl

1 Center of Education and Training, University Medical Center Utrecht, Utrecht University, Utrecht, The Netherlands

2 Center for Molecular Medicine, University Center Utrecht, Utrecht University, Utrecht, The Netherlands

1 Introduction

The increasing size and complexity of biomedical data bring great opportunities, but they also require advances in scientific methods to extract knowledge from these data. In addition, it is becoming clear that the development of the field of biomedical data science is accompanied by an increase in scientific flaws and questionable research practices (Altman & Levitt, 2018; Benchimol et al., 2015; Elliott et al., 2015; Mayer, 2018; Nie et al., 2018; Valcu & Valcu, 2011; Vaught et al., 2017). Biomedical *education* should also respond to these developments. To that end, several players in the field of biomedical education have been working on defining biomedical data science learning goals for university students in biomedical sciences (Attwood et al., 2015; Corpas et al., 2015; Dinsdale et al., 2015; www.elixer-europe.org). In some central publications, members of these networks have formulated learning goals specific to biomedical data science (or bioinformatics) (Tractenberg, 2017; Tractenberg et al., 2019; Wilson Sayres et al., 2018). However, these lists are extensive and contain many specific actions or behaviors that are required of students. Moreover, some learning goals require many hours of training for students to reach required levels (e.g., programming and using bioinformatics software). We argue that instructional practices in biomedical data science education should not focus on teaching practical skills but on the general skills, mindsets, and views underlying the formulated learning goals. This fosters a scientific way of thinking and acting which is needed to adequately conduct scientific research.

At the same time, players in the field of philosophy of science have been working on defining how they should train (biomedical) scientists in philosophy of science (Boniolo & Campaner, 2020; Grüne-Yanoff, 2014). They provide strong rationales for the increased need of teaching philosophy of science to scientists and for a thorough revision of what is taught in these courses. However, they also discuss current challenges to teaching philosophy of science to scientists. The major challenges discussed include that students have no training in the humanities and do not see the importance of notions fundamental to philosophical work such as argumentation (Grüne-Yanoff, 2014). In addition, students have no intrinsic motivation to take a philosophy course (Grüne-Yanoff, 2014) and often only become interested in philosophical issues in the later stages of their careers (Boniolo & Campaner, 2020). Further, many students have not yet engaged in scientific research and lack knowledge of basic scientific research practices such as designing an experiment (Grüne-Yanoff, 2014). In addition, the student body of such courses will be large which makes the degree of interaction low (Grüne-Yanoff, 2014). Lastly, philosophers first need to know what scientists really need from philosophy before they can revise their training (Boniolo & Campaner, 2020). Both articles offer some solutions to these challenges such as teaching scientists the philosophical work of analyzing scientific procedures and of evaluating scientific knowledge production rather than teaching them the main epistemological and ethical theories (Boniolo & Campaner, 2020; Grüne-Yanoff, 2014). These are important revisions of standard philosophy of science curricula. Therefore, we discuss the value of these and their other recommendations for biomedical data science education in this article. However, they are both limited by the presupposition that philosophy of science should be taught in a dedicated course. This limits the so desired applicability of the philosophy of science training to students' own practices. Even when new philosophical skills are applied to relevant examples during such courses, students are unlikely to see these philosophical skills as part of being a biomedical scientist. They will be more likely to see them as skills belonging to another discipline. Therefore, we cannot expect them to automatically transfer these skills to their own research practice as well.

In summary, we observe two concurrent challenges. The first is that biomedical sciences need scientists that can conduct rigorous research even with increasingly complex and large datasets without getting lost in training in many technical skills and methodologies. The second is that philosophy of science needs to be taught in such a way that it can be applied to everyday research practice. In this article, we investigate the possibility of approaching these challenges together.

2 Our Approach

As biomedical sciences teachers, our original research focus was the challenge of biomedical data science education. So, we started with analyzing recent literature about how to prepare students for conducting research with increasingly large and complex datasets. We found that others proposed learning goals for biomedical data science education to address the changing requirements placed on biomedical scientists. We identified the articles of Wilson Sayres et al., (2018) and (Tractenberg 2017; Tractenberg et al., 2019) as key publications. We analyzed their learning goals to assess their approach to this challenge. As a result of our analysis, which we detail in the remainder of this article, we identified five important aspects of conducting research (such as "make and justify decisions") in these long lists of learning goals. In addition, we came to realize that biomedical data science skills can be trained on, what we describe as, four dimensions of learning. We established that these dimensions increase from practical aspects to a more conceptual understanding. In "Sect. 4," we provide descriptions of the five key aspects of conducting research and four dimensions of learning, as well as explanations of how we identified them. However, for the purpose of explaining our approach, we already highlight one of our observations. During our analysis, we realized the highest dimension of learning had to do with understanding the *nature* of science. However, the authors of the learning goals (Tractenberg, 2017; Tractenberg et al., 2019; Wilson Sayres et al., 2018) do not explicitly address the nature of science in their learning goals. Although understanding the nature of science has been and remains an important objective of science education, it does not seem to be on the forefront of educational strategies for teaching and learning *biomedical data science* in university degree programs. Therefore, we investigated strategies proposed elsewhere for teaching students about the nature and philosophy of science (e.g., Abd-El-Khalick, 2012; Boniolo & Campaner, 2020; Grüne-Yanoff, 2014; Lederman, 2007) and compared them with the learning goals of biomedical data science education. Through our comparative reading, we explored the possibility of using philosophy of science training approaches to prepare students for handling increasingly large and complex biomedical datasets.

In addition to identifying key aspects of conducting research and dimensions of learning, we address some other facets of the discussion of the role of philosophy of science for teaching biomedical data science and vice versa:

- We argue that current and proposed educational approaches for biomedical data science do not sufficiently attend to the higher dimensions.
- We argue that proposed educational approaches for philosophy of science do attend to these dimensions but risk losing the connection with the lower dimensions and therefore with everyday research practice.

- We propose design principles for an integrated educational approach that simultaneously attends to all four dimensions of learning to conduct research.

To structure our discussion, we start with a brief account of the challenges faced by the field of biomedical data science and of the current and proposed educational approaches to address these challenges. Then, we identify the first aspect of conducting scientific research which leads us to the four dimensions of learning. Next, we successively discuss each of the other aspects of conducting research that we identified. For each key aspect, we explain how we identified the aspect, and we discuss how philosophy of science and biomedical data science training can be integrated to train biomedical data science skills on all four dimensions by explicitly attending to the second to fifth key aspects of conducting research (summarized in Table 2).

3 Developments in Biomedical Data Science and Biomedical Data Science Education

Biomedical science is a broad field in which many different research techniques are used. Biomedical studies vary from visualizing a few molecules in a few cells or a single tissue sample to genotyping or sequencing DNA of thousands of people. There is a broad range of laboratory techniques a biomedical researcher can choose from to, for example, sequence DNA, to separate, identify, quantify, or visualize molecules, to determine molecular structures, to characterize cell processes, or to determine effects of drugs or other compounds. In addition, many of these techniques are difficult to master and technical mistakes are easily made. Further, the resulting data have strongly increased in size and complexity.

The reason for increasing size and complexity of data is that the amount of data on human health and disease we *can* acquire and process in a short time has increased strongly over the last years. This so-called data deluge causes rapid advances in knowledge in the field of biomedical sciences. Large datasets, for example, provide the opportunity to explore many hypotheses at the same time to quickly narrow down to hypotheses that are useful to test in subsequent research. However, the advances also reveal that life, health, and disease are a lot more complex than we previously believed. Hence, development of scientific methods for extracting knowledge about life, health, and disease from these data is required (Braillard, 2013). Development is needed in the design and execution of acquisition, annotation, organization, analysis, and interpretation of biomedical data. Building on previous work, we use the term "biomedical data science" to refer to these activities (Altman & Levitt, 2018).[1]

It is also becoming clear that the development of the field of biomedical data science is accompanied by an increase in scientific flaws and questionable research practices. There are regular concerns with validity of data, methods, or analysis practices (Elliott et al., 2015; Mayer, 2018; Valcu & Valcu, 2011; Vaught et al., 2017). In addition, the reporting on data and important information to put results in a larger context is often insufficient for subsequent validation (Altman & Levitt, 2018; Benchimol et al., 2015; Nie et al., 2018). Especially with new types of inquiry and new types of data it is challenging to ensure the

[1] Others have used the term "bioinformatics" to describe these activities (Attwood et al., 2019). This is also the term that is used in the discussed learning goals. However, we believe that, as Altman and Levitt (2018) argue, the term "biomedical data science" is more applicable.

validity and reliability of research and analytical methods for the specific research question. This requires profound understanding of methodology and statistics and adequate reporting of research.

New generations of scientists are an important factor in improvement of research practice. From the bottom up practices can be reformed. Therefore, we should improve biomedical data science *education* to bring about the change we are looking for here. Although not the focus of this article, it is good to note that the formal education of current researchers was not tailored to the increasingly complex data and analysis methods that are required today. Therefore, it would also be useful to reform continued professional development of biomedical scientists, for example, in seminars and conferences. However, here we focus on university degree programs.

Throughout the world, players in the field of biomedical education have been defining the knowledge, skills, and abilities that should be taught to young biomedical scientists to bring about the desired change. Members of some key global networks have formulated lists of required competencies and the knowledge, skills, and abilities that lead to these competencies (Tractenberg, 2017; Tractenberg et al., 2019; Wilson Sayres et al., 2018). In short, Wilson Sayres et al. (2018) formulated nine core competencies in biomedical data science for biomedical undergraduate students. These are quite general competencies which each imply multiple skills, behaviors, and attitudes that are required to show the corresponding competency. Therefore, Tractenberg et al. (2019) developed a mastery rubric containing performance levels descriptions of what students at five different levels of expertise in bioinformatics should achieve in terms of knowledge, skills, and abilities. They describe around fifty skills, behaviors, and abilities per level, divided over twelve categories. They developed a similar mastery rubric for statistical competencies (Tractenberg, 2017), which, in our definition, fall under biomedical data science competencies as well. This rubric contains some redundant and some unique categories compared to the bioinformatics rubric. Here, we use both the bioinformatics and the statistical literacy mastery rubric. As a side note, we recognize the differences between competencies, knowledge, skills, and abilities, and performance level descriptors, but refer to the articles of Tractenberg and colleagues for a detailed explanation of these differences (Tractenberg, 2017; Tractenberg et al., 2019). For ease of reading, we will collectively refer to these with the term (biomedical data science) "learning goals."

As Tractenberg et al. (2019) point out, their learning goals characterize phases of a scientific process (described as "the scientific method" by Tractenberg, e.g., 2019, p. 4). Students need to be able to define a problem, design research, acquire and process data, analyze processed data, interpret their analyses, and communicate those results and interpretations. For each of these steps, (Tractenberg 2017; Tractenberg et al. 2019) describe several specific learning goals. A major strength of these learning goals is that they are not content-specific. They focus more on general actions and considerations (e.g., "can initiate a search for data & will ask if uncertain about relevance for any given problem" (Tractenberg et al., 2019, p. 11)) than on domain-specific procedures (e.g., "navigate and retrieve data from a genome browser" (Wilson Sayres et al., 2018, p. 14)). They point out that the content-independence of their rubric is beneficial because of its applicability to various sub-disciplines and constantly changing research methods and technology (Tractenberg et al., 2019). In addition to the advantage of the broad applicability for future students, we argue, it is also advantageous for current biomedical data science students. They will be the ones using or developing new research methods and technology which requires different skills than the skills needed to use current methods. We will elaborate on this line of thought in the remainder of the article.

4 Five Key Aspects of Conducting Scientific Research and How to Train Them Through Philosophy of Science

4.1 Key Aspect 1: Understand Data, Research Methods, and Statistical Methods, and Use Them

The first eight out of nine (domain-specific) core competencies describe different aspects of understanding and using biomedical data and methods (Wilson Sayres et al., 2018). In short, students should know what types of data and software exist (core competency 1) and have a basic understanding of key computational concepts such as the organization of biomedical databases (2). In addition, they should be able to apply statistical concepts used in bioinformatics (3); use bioinformatics tools to examine problems (4); find, retrieve, and organize various types of biological data (5); explore and model biological interactions (6); and use command-line or shell scripting (7). Lastly, students should understand and use different biological data types and recognize that all experimental data is subject to error and recognize the need to verify the reproducibility of data (8) (Wilson Sayres et al., 2018). Together, these competencies could be summarized as “understand data, research methods, and statistical methods, and use them.” This is the first key aspect of conducting research we identify in the biomedical data science learning goals.

These competencies mainly describe the practical tasks biomedical data scientists are facing or the methods they need to be able to use. Exceptions are the formulation of the eighth competency because it includes recognition of factors influencing results of data analysis as well, and the ninth competency discussed later in this article. Nonetheless, these proposed new competencies are mainly aimed at an expansion of the biomedical scientist’s toolkit for acquiring and analyzing biomedical data to enable analysis of larger and more complex datasets. This focus on new practical requirements is also reflected throughout most other literature about biomedical data science education which we, albeit not systematically, reviewed (e.g., Dill-McFarland et al., 2021; Hicks & Irizarry, 2018; Horton, 2015; Kleinschmit et al., 2019; Madlung, 2018; Olimpo et al., 2018; Rosenwald et al., 2016). Thus, they mostly focus on our first identified key aspect of conducting biomedical data science research. Tractenberg et al. (2019) already recognized that the learning goals they describe themselves are of a different type than these competencies described by Wilson Sayres et al., (2018). They contend that competencies are what individuals can do when they bring their knowledge, skills, and abilities together. They emphasize the need for defining the route to achieving these competencies instead of just describing the endpoints. And they see the solution in determining different levels of knowledge, skills, and abilities which form a developmental trajectory toward achieving competencies (Tractenberg et al., 2019). We agree that this is an important shift of focus for development of biomedical data science education. However, we argue that the shift they make encompasses more than a shift from defining endpoints of educational programs to defining the route toward these endpoints.

4.1.1 Interlude: Four Dimensions of Learning

When we compared the core competencies with the learning goals in the mastery rubrics, we noticed that they describe different *dimensions* of biomedical data science skills. They describe the same research skills, such as experimental design and interpretation of research results, but aim at different dimensions of these skills. The core competencies (Wilson Sayres et al., 2018) are mainly focused on teaching students how to apply different techniques in biomedical data science and to understand how these techniques work. The mastery rubrics (Tractenberg, 2017; Tractenberg et al., 2019), on the other hand, are mainly focused on learning to design research approaches and make choices during different phases of a research project. This is the difference between how you do something (core competencies) and what you choose to do (mastery rubrics). Or, in other words, it is the difference between knowing how to push buttons and choosing which buttons to push. When analyzing data, for example, a researcher needs to choose a method of analysis (e.g., multiple linear regression) and needs to conduct the analysis (e.g., write the script). While the common skill is conducting a statistical analysis, the former is a higher dimension of that skill than the latter.

Then there is also a distinction between choosing what to do based on instructions or checklists and understanding what those choices are based on. The latter forms a third dimension of the same research skill. Tractenberg et al. (2019) touch upon this difference when they write: "Like competencies, [knowledge, skills, and abilities] can be highly complex; by contrast, however, [knowledge, skills, and abilities] are general-i.e., the same [knowledge, skills, and abilities] can be deployed differently to support different task-specific competencies" (p. 3, punctuation in original). However, we contend that the difference is not necessarily in the often-made distinction between general and task- or discipline-specific skills but in the dimension of application of the same skill. Both discipline-specific and general skills can be demonstrated at different dimensions of application and understanding. Training students on the first dimension, learning how to apply research techniques, focuses on the first aspect of conducting research. The core competencies, like many other biomedical data science education approaches, mainly focus on this first aspect by expanding the student's toolkit with new data analysis techniques and software. However, training students in the use of all these new tools will require a great deal of precious time spent on data science skills. Furthermore, in our experience, students tend to focus on technical skills (e.g., getting their script to work), which makes it difficult for teachers to bring across a general understanding and the nuances of data analysis at the same time.

However, with the current developments in biomedical sciences, there is not only a need for new tools but also a need for more rigorous use of these and the old tools. Scientific flaws and questionable research practices are not counteracted by use of new tools. Rather, they require certain ways of thinking and acting that transcend practical actions (i.e., transcending the first dimension). Therefore, we should focus on teaching students which research design choices to make (second dimension) and, even more important, to understand the rationale for these choices (third dimension). This second and third dimension of learning we did identify in the mastery rubrics (Tractenberg, 2017; Tractenberg et al., 2019). Instead of focusing on the practical tools and workflows, they focus on other aspects of conducting scientific research. We discuss our identification of these aspects in the remainder of this section. We explain that by choosing to focus on these aspects of conducting scientific research, educators are implicitly steered to educational approaches that align with the second and third dimension of learning. We argue that learning to attend to these other aspects and dimensions will help students with learning to understand data, research methods, and statistical methods, and using them *appropriately*. Furthermore, we add an important fourth dimension of learning to conduct scientific research which we do not recognize in the proposed learning goals. This fourth dimension is understanding the nature of science and this understanding can help students with understanding the rationale

Table 1 Four dimensions of learning to conduct scientific research

Dimension	Examples
1. Applying techniques	Writing a script, using software, performing a western blot, cell culture, etc
2. Making choices in research design and implementation	Selecting data, choosing a statistical test, reporting research according to provided standards, etc
3. Understanding the rationale behind choices	Awareness that not all data are relevant, understanding the meaning of a *p*-value, understanding what data to present in publications, etc
4. Understanding the nature of science	Understanding that research choices are not always about right and wrong, understand that each choice has consequences for the value and quality of the results, understand that data do not speak for themselves but require human interpretation, etc

for choices in research design and implementation. The four dimensions are summarized and exemplified in Table 1.

So, the first aspect of conducting scientific research is understanding data, research methods, and statistical methods, and using them. By focusing on this aspect of conducting research, biomedical data science educators tend to focus on the first dimension of learning by teaching how to apply biomedical data science techniques. Instead, we should focus on four other aspects of conducting scientific research that have to do with the higher dimensions, with making choices in research design and implementation, and understanding the rationale behind choices. Furthermore, using a fourth dimension of learning, understanding the nature of science, can provide a means to achieving a way of thinking and acting that transcends practical actions and fosters better scientific practice. Throughout the remainder of the article, we elaborate upon how we recognize the three identified dimensions of learning in the learning goals and how an addition of the fourth dimension can benefit learning. To this end, we discuss the second to fifth aspect of conducting research we identified in the learning goals and how these can be taught in such a way as to include the higher dimensions of learning.

4.2 Key Aspect 2: Make and Justify Decisions

As we have argued, the mastery rubrics for bioinformatics and statistical literacy (Tractenberg, 2017; Tractenberg et al., 2019) do focus more on certain ways of thinking during research design and implementation than the core competencies (Wilson Sayres et al., 2018) do. However, the list of learning goals is so long and specifically applied to each phase of research that it has two drawbacks. The first drawback is that the list is so long that it is no longer feasible to cover all choices and considerations in an educational program. (Tractenberg 2017; Tractenberg et al., 2019) are very thorough in describing the many choices that must be made in research design and implementation (hereafter "choices" or "decisions"). For example, in the learning goals in the category "identify data that are relevant to the problem" we find that "[students should learn] how to identify, & evaluate strengths & weaknesses of, data sources, to determine whether a given data-set or -type is relevant for a given problem" (Tractenberg et al., 2019, p. 11). And, in the category "identify & use appropriate analytical methods" we find that "[students should learn] to evaluate/rank & justify alternative methods in terms of general features of their efficiency & relevance for the given research problem" (Ibid.). For other phases of a research process similar choices are described in the learning goals (Tractenberg, 2017; Tractenberg et al., 2019). In addition, they also describe many *considerations* that are important in making those choices. This results in approximately forty specific choices and considerations. In our view, it will be difficult to pay explicit attention to all forty in a single bachelor's program. The second drawback is that there is a risk that by learning all these specific considerations, students treat those considerations more as a step-by-step guide than as a way of thinking and acting. As a result, students mainly learn to work at the level of the second dimension (making choices based on instruction) instead of developing an understanding of rationales for choices (third dimension).

However, our identification of these choices and considerations as recurring elements in the learning goals leads us to the identification of a second key aspect of conducting research. That is, making and justifying decisions throughout the research process. This is an aspect of conducting research that is implied in the learning goals but is not made explicit by the authors. It can be beneficial to bring this aspect to the fore. This can be done by using the specific choices and considerations mentioned as examples of the process of making and justifying choices instead of treating them as separate knowledge, skills, and abilities. Then it is no longer necessary to have students encounter all possible moments of choice and go into the details of all these choices during their education. Yet, as we explain below, it does provide them with the tools to handle various choices appropriately in the future. This approach focuses more on teaching a way of thinking and acting (third and fourth dimension of learning) than on teaching specific considerations and actions (second dimension).

4.2.1 Train Explication of Considerations for Scientific Decisions

There are some frequently recurring concepts that underlie the specific considerations detailed in the learning goals that can help students to explicate these specific considerations for scientific decisions themselves. The recurring concepts are relevance, assumptions, uncertainties, reproducibility, and rigor (Tractenberg, 2017; Tractenberg et al., 2019). These are important concepts to keep in mind during decision-making. Therefore, we will discuss them more elaborately than the authors do. The degree to which data and methods are useful to the research problem (i.e., relevance) impacts the applicability of the resulting knowledge to the solution of the research problem. Similarly, the assumptions behind hypotheses and methods, and other factors within the research design affect the degree of certainty that can be awarded to results and conclusions and the circumstances under which the conclusions are thought to hold. Further, the degree to which a scientific method can be or was strictly applied to make scientific processes well-controlled and to decrease bias (i.e., rigor) and whether the results can be supported by different researchers using the same experimental setup (i.e., reproducibility[2]), are factors that guide decision-making of scientists as well.

These factors (i.e., relevance, assumptions, uncertainties, reproducibility, and rigor) could be presented by teachers during lectures or working groups. However, this risks

[2] We acknowledge that there are distinct types of reproducibility and although this distinction is important, we did not think it relevant to describe it in this article. Discussion of forms of reproducibility in biomedical research can be found elsewhere (Montgomery, 2019).

again that students treat it as a step-by-step guide that can be followed blindly. Instead, we propose to guide students in producing these considerations themselves. Like philosophers of science, they should reflect themselves on decision-making in scientific studies and make explicit what concepts and principles guide a scientist's decisions. Here, we see a difference between making choices based on instructions (i.e., the second dimension of learning) and understanding the underlying rationales that then lead to the consideration of the specific factors and principles (i.e., the third dimension). To help students discover these factors themselves, we can borrow from the procedures of philosophy of science.

Grüne-Yanoff (2014), for example, suggests that the hypothetico-deductive model, the falsification concept, or the model of inference to the best explanation can help analyze scientific methodology. He argues that these tools that are taught in standard philosophy of science courses remain relevant (Grüne-Yanoff, 2014). However, he does not go into detail about the role these tools can play in instruction. To make this recommendation more concrete, we suggest that students could learn about these distinct types of inference and that these all have their own scope and limitations. But more importantly, they could then learn to recognize these types of inference in research designs when they do their own research or when they discuss research designs during biomedical courses. The ability to abstract research designs to general principles of argumentation, and knowledge about these principles could help students to identify the scope and limitations of different methods and to produce rigorous and reproducible research. To that end, students should be supervised in these philosophical exercises at times when research designs are also covered from a disciplinary or data science perspective, in biomedical courses.

In addition, what characterizes a philosopher's working method is the ability to take apart a process or an argument to identify its underlying doctrine. With the help of philosophy of science, educators can teach students how they can identify the principles, beliefs, and values that guide their own and another's decision-making process. And to recognize the underlying concepts of research and analysis methods. This taking apart and abstraction can help students to make and justify their decisions.

These philosophical insights are especially, although not exclusively, useful in the field of biomedical data science. The increasing size and complexity of data are accompanied by an ever-increasing number of data analysis methods. Students cannot master all these methods in the limited amount of time they have. However, we can provide students with tools that prepare them for learning to master new data analysis methods in later career stages. That is, by teaching them the philosophical procedures to take apart research and analysis methods and to make (considerations for) decisions explicit. In addition, this experience can help researchers in developing new methods for new challenges in this rapidly advancing field.

However, it is good to note that philosophy of science training should also focus on ways of thinking about scientific research and ways of approaching research design and implementation rather than learning to use specific tools. Tools such as the hypothetico-deductive model are useful for understanding certain research approaches, but they are not applicable to all of science. Especially studies working with large and complex datasets are using new types of research design and analysis. Therefore, the tools of traditional philosophy of science are not sufficient for dealing with big data research (e.g., Ernst, 2009). Despite, or perhaps even more so *because of*, these developments it can be worthwhile to teach students these philosophical ways of thinking and acting and how to explicate considerations for scientific decisions.

4.2.2 Train Reflection on the Principles and Concepts of Probability and Statistics

Abstracting specific methods to general principles can also help students to make appropriate decisions in *statistical* methods and outcome measures, an essential part of data science education. In the learning goals, we recognize the importance of learning to evaluate the measurement properties of variables. Understanding how these variable properties determine the choice of statistical methods is essential for choosing appropriate methods (Tractenberg, 2018). In addition, the learning goals define that students should understand the roles of covariates, what the p-value represents, and how to apply false discovery rate controls and why (Tractenberg, 2018). Thus, the learning goals essentially say that students learn to make appropriate statistical decisions by recognizing that they should consider the features of the data. As one of our reviewers kindly pointed out, these learning goals focus mainly on null hypothesis significance testing (NHST). Although NHST remains the most frequently used statistical approach in biomedical sciences, it is crucial to consider the trend toward other statistical approaches, like model selection or using Bayes Factors. In this light, we first discuss how achieving the NHST-oriented learning goals can be facilitated by philosophy of science approaches and then discuss the consequences for statistical approaches not mentioned in the learning goals.

Again, learning how decisions in specific methods are related to general concepts in scientific methodology could help students to recognize these concepts in their own research. The general concepts that are applicable here are probability and uncertainty. It could be argued that all statistical tools measure the uncertainty of data and results (Lindley, 2000). Therefore, it could be beneficial for students to reflect on the role of probability and uncertainty in science on a more abstract level. One aspect of this reflection could be on the following distinction between probability and significance level, as described by Lindley (2000), which is relevant in the most frequently used method of NHST. A scientist starts with a hypothesis about an *uncertain* process. They acquire data to remove or reduce this uncertainty, and they use the probability of the acquired *data* (or more extreme data) given that the null hypothesis is true (i.e., significance level), instead of the probability of the *hypothesis* given that the data are true (i.e., probability). In other words, in NHST, scientists assume that the hypothesis is true to determine how probable it is that the data in question have been found. While, in reality, the truth of the hypothesis is under discussion (Lindley, 2000). We argue that students should realize this is the case and could reflect on its consequences. This constitutes the important fourth dimension of learning to conduct scientific research (Table 1). It is in the *nature* of null hypothesis testing that you determine the probability of the data given the null hypothesis rather than determining the probability of a research hypothesis. Understanding this nature can be key to drawing appropriate conclusions, yet it is often not at the forefront of data science education approaches.

This way of thinking is extra important because there is a trend toward using other approaches than NHST in scientific research and statistics education. These include contrast testing (replacing the classical alternative hypothesis with a pre-specified contrast), equivalence testing (testing the existence of the null hypothesis allowing for a minimum effect), inequality constrained hypothesis testing (applying inequality constraints between the parameters of interest), Bayesian statistics (starting with a prior view on the probability), and model selection approaches such as structural equation modeling. Even though most biomedical research approaches still exclusively use NHST, these trends toward other methods enhance the need for teaching students to reflect on the use of probability in science on a more abstract level. All methods result in estimations of how well a theory represents a phenomenon and the way they obtain these measures of uncertainty affects the conclusions drawn from the analysis. Firstly, it is paramount that scientists realize *that* they do have a choice of method. Secondly, they should learn about the consequences of each choice for the interpretation of their results (see the next subsection). We contend that we should start by teaching students the first lesson even though they might not yet be able to oversee the specific consequences. We might even need to repeat this *ad nauseam* throughout all our lessons about interpreting results: all statistical results are estimations of uncertainty in terms of probability and the value of these probabilities depends on the decisions made throughout data collection and analysis and on the intended use of the results.

A concurrent trend in biomedical sciences is the increasing use of approaches that depend less on hypotheses and theories and that are often more data driven. Examples include advanced data visualization techniques and clustering techniques to reduce high-dimensional data. The output of these approaches is mostly the model itself and not necessarily theory building. For example, in machine learning techniques, a model is tested and validated on large datasets, and this validated model is the product of the study rather than an increased understanding of the parameters measured in the datasets. The traditional tools of philosophy of science might not be applicable to these approaches. However, the way of thinking and acting of a philosopher, which we promote here, can be beneficial for students to make more appropriate use of these approaches. Furthermore, it shows students again that there are many methods in science and that there is not one "correct" way of proceeding.

Another aspect of the reflection on probability and uncertainty could be on the assumptions behind statistical concepts (e.g., *p*-value, significant, confounder, effect size). Boniolo and Campaner (2020) suggest that reflection on statistical concepts can provide understanding of the assumptions behind these concepts and on their different interpretations and chosen meanings (Boniolo & Campaner, 2020). For this matter, we add that philosophers of science can share their insights in the use of statistical concepts with students to open their minds to the idea that there *are* different meanings of statistical concepts. This could then incite caution in their own use of the concepts and in the interpretation of their use by others. These reflections strongly relate to the human-constructed and socio-cultural embedded nature of science. Therefore, discussing these aspects of the nature of science, using the negotiation of the meaning of statistical concepts as example, can also be helpful to increase student understanding. This is another example of learning on the fourth dimension. Students are not only helped to make appropriate choices during research by understanding the rationale for these choices. In addition, it can be helpful for them to understand the nature of scientific research and, for example, that our conceptions affect how we interpret scientific terms such as "probability" and "significance."

Of course, we are not the first to mention that reflection on the human-constructed and socio-cultural embedded nature of science should be included in science education. With this we touch upon yet another research field, that of nature of science education. Nature of science scholars, for example, also advocate teaching the social character of science (Abd-El-Khalick, 2012; McComas, 2020). They propose to propagate the view that science uses an intersubjective, collaborative approach which minimizes researcher bias and subjectivity (McComas, 2020). These and other aspects of the nature of science are discussed by many scholars of nature of science education. Generally, these scholars discern seven consensus views of the nature of science that students should learn during their formal education (Lederman, 2007): scientific knowledge is tentative, empirically based, subjective, human-constructed, and socially and culturally embedded. In addition, there is a distinction between observations and inference and between scientific laws and theories (Lederman, 2007). This listing is neither exclusive nor exhaustive but forms the basis of nature of science education. Since these aspects of the nature of science are intricately linked to what we discuss, we also discuss how some of them can be attended to in our proposed training.

From this point of view, we should also look back at the first key aspect that we identified (understanding and using data and methods). The tools and methods discussed in the learning goals are grounded in the empirical nature of science. Therefore, we implicitly teach that scientific research has an empirical nature. However, making this aspect of the nature of science explicit could be a first step toward discussing the nature of science in a broader sense with students. A logical next step would be to examine the role of the scientist in the creation of scientific knowledge. Discussing the human-constructed and inferential nature of science opens up the conversation about the many choices that are made by scientists throughout a research project. This way, students can recognize that understanding and using data and methods is only one aspect of conducting scientific research and that the other aspects (as discussed throughout this article) require a scientist who actively draws inferences and chooses between reasonable alternatives.

4.3 Key Aspect 3: Explain How Decisions Affect Results

In the last section, we have already pointed out that it is important that students learn to make and justify decisions because a researcher does make many choices throughout the research process and their decisions affect results. However, learning that decisions do affect results is so important that it calls for separate attention. We also recognize this importance throughout the learning goals (Tractenberg, 2017; Tractenberg et al., 2019). For example, in the category "Define a problem based on a critical review of existing knowledge" the mastery rubric states: "[An undergraduate] … in guided critical reviews, [is] learning to recognize that design features & evidence base are important to drawing conclusions" (Tractenberg et al., 2019, p. 10). And, in the category "Experimental Design" it states:

> [An undergraduate is] developing the understanding that weak experimental design yields weak data & weak results. Needs assistance in conceptualizing covariates & their potential roles in the planned analyses. Beginning to recognize that, & can explain why, just one study is usually insufficient to answer a given research problems/solve biological problems adequately. (Tractenberg et al., 2019, p. 10)

Again, the learning goals mention several times that students should understand that specific aspects of scientific research (such as research design) and decisions made (such as doing unplanned analyses) affect the results. So, here we identify the third key aspect of conducting research of "explaining how decisions affect results."

4.3.1 Train Reflection on Effects of Epistemic Aspects of Decision-Making

This aspect shows as well that attending to the higher dimensions of learning to conduct research is important and that we should train students to reflect on the effects of epistemic aspects of decision-making. An example of an assignment in undergraduate programs where students can learn to explain how decisions affect results is writing a research proposal. In such an assignment, students need to learn how to work with a research proposal format and need to learn what type of information they need to provide in which part of the proposal. This is the very practical, first dimension of the assignment. Then, there are of course many choices to be made. Undergraduate students of course require guidance in this regard, since they do not yet have an overview of possible research methods nor enough experience in the field. However, we can provide diverse types of guidance. In these types of assignments, students often look for the "correct" answer or for the "right" research design for their research question. In other words, they might seek to learn on the second dimension. However, this assignment provides an opportunity to foster the idea that there is not one correct or best method for each research question. In science as we practice it today, we have, as Boniolo and Campaner (2020) also argue, a certain way of proceeding, the current scientific methodology, which is refined over the course of millennia. In addition to this general scientific paradigm, every scientist operates in a disciplinary paradigm with its own methodology (Boniolo & Campaner, 2020). However, this methodology has its assumptions and even when followed rigorously, it cannot guarantee the validity of the results. So, we argue that we need to get students into the habit of considering the underlying assumptions of the methods they want to propose in their research proposal. And to consider the implications of those assumptions for how they might interpret the results of the proposed study.

4.3.2 Train Discussion of How Scientific Research Contributes to Theory Building

Readers might assume it is evident for students that scientific methodology cannot guarantee the validity of results and that students will not hold the belief that science provides certain knowledge about the real world. However, although only a small number of studies have explored scientists' views about the tentativeness of scientific knowledge (Yucel, 2018), some life scientists did express the view that science attains certain knowledge (Schwartz & Lederman, 2008). In addition, the way biomedical studies are reported (see "Sect. 3") seems to suggest that some (or even many) biomedical scientists hold false beliefs about the truth value of their conclusions. If students are convinced that good scientific research *should* provide definitive evidence for a theory, that is potentially dangerous. It could, for example, lead to ignoring the consequences of choices in the research design for the epistemic status of the conclusions. Appreciating the value of taking *a* step and providing evidence that supports or does not support a hypothesis *to a certain extent* will do more justice to the reality of research. We should help students to realize that scientific research provides arguments to make a theory more (or less) supported. Then we could, as Abd-El-Khalick (2012) also proposes as an objective of high school (nature of) science education, shift their focus from proving or disproving theories to arguing the extent to which a theory can explain or predict phenomena and under which circumstances. So, we should discuss with students how scientific research contributes to theory building.

4.4 Key Aspect 4: Balance Epistemic, Ethical, and Societal Responsibilities

In the biomedical data science learning goals, we also found goals that are focused on the ethics of research. The mastery rubric (Tractenberg et al., 2019) describes that students should know that it is important to learn to recognize scientific misconduct and unethical practice, and how to act ethically. For example, by stating that "[An undergraduate is] learning the principles of ethical professional & scientific conduct. Seeks guidance to strengthen applications of these principles in own practice. Learning how to respond to unethical practice" (Tractenberg et al., 2019, p. 9). In addition, they included "key attributes of ethical science in all of the [learning goals] relating to transparency, rigor, and reproducibility" (Tractenberg et al., 2019, p. 8). An example of this can be found in a learning goal in the category "Define a problem based on a critical review of existing knowledge": "[An undergraduate] recognizes the role of uncertainty in research, & that reproducibility & potential bias should be considered for every result" (Tractenberg et al., 2019, p. 10). The core biomedical data science competencies (Wilson Sayres et al., 2018) add interpretation of the ethical, legal, medical, and societal implications (ELMSI) of biological data as their ninth and last competency. This learning goal is focused on the implications of the types of data biomedical sciences are dealing with (e.g., the availability of genomic data) and the procedures that protect against falsification or manipulation of biomedical data (Wilson Sayres et al., 2018). So, the learning goals describe that students should learn how to act ethically and how to discuss the implications of biomedical data science *for people*. They contend that students need to learn how to act ethically responsibly in conducting scientific research and to learn what procedures we have in place to hold scientists to these ethical principles. So, an additional aspect of conducting scientific research we can identify in the learning goals is to act upon ethical and societal responsibilities. What is not included in the learning goals is how ethical conduct and considering the implications of scientific research can affect *scientific research itself* nor to what extent it does so. Scientists need to balance their ethical and societal responsibilities with their epistemic responsibilities. Therefore, we suggest adding epistemic responsibilities to the fourth aspect of conducting research. Further, we propose a focus on learning to balance these epistemic and non-epistemic considerations rather than just learning the principles of ethical conduct and learning to discuss societal implications of research. Therefore, our fourth key aspect of conducting research is "balance epistemic, ethical, and societal responsibilities."

4.4.1 Train Identification of Underlying Values of Epistemic and Non-epistemic Considerations

To that end, it is first of all important that students learn to recognize the deeper values underlying the epistemic and non-epistemic considerations. For example, reproducibility, rigor, and transparency are valuable because they enable separating science from other kinds of knowledge and preventing negative impacts of scientific results on people (Boniolo & Campaner, 2020). So, philosophy of science points to a deeper layer of values underlying these three. We argue that helping students gain an understanding of the general principles of science and the deeper layer of underlying values could help them understand the influence of science on society and vice versa. These insights could subsequently contribute to greater motivation to include both epistemic and non-epistemic considerations in research decisions. And to provide means for balancing them responsibly.[3]

For learning to balance epistemic, ethical, and societal responsibilities, it would be most valuable if students learn to identify values and norms themselves instead of presenting them with commonly held values and contemporary rules and guidelines that they should

[3] Naturally, these educational approaches are closely related to ethics education. However, further elaboration on ethics education is beyond the scope of this article and can be found elsewhere (Gerrits et al., 2021).

adhere to. Therefore, we need to show students what questions to ask about ethical conduct and societal implications and *how* to answer them. Questions should, for example, include: What values could potentially conflict here? Which stakeholders are involved? What are the needs and wishes of these stakeholders? Are these in conflict between stakeholders? How do I protect the anonymity of participants? Am I aiming for anonymity or pseudonymity? By learning to ask questions rather than to adhere to guidelines, students are more actively involved in the act of balancing different considerations. This creates two advantages. The first advantage is that students are actually made responsible for considering all important ethical and societal aspects of their scientific research practice. The second advantage is that it, again, fosters the idea that scientists do need to balance different responsibilities and that they therefore need to actively choose between reasonable alternatives rather than simply adhere to rules and guidelines and make the "right" choice.

4.4.2 Train Reflection on Effects of Non-epistemic Aspects of Decision-Making

Which non-epistemic aspects should researchers consider then? Grüne-Yanoff (2014) also argues that there are non-epistemic considerations underlying decisions in the scientific process and that these should be justified and evaluated. His considerations include scientific honesty,[4] ethical conduct of experiments, considering consequences of one's research, the role of policy making, and the effects of a scientist having an expert role in the political domain (Grüne-Yanoff, 2014). Scientific dishonesty, he implies, is distorting the research process, for example, by misreporting empirical findings, plagiarism, and denigrating the involvement of co-researchers. Ethical conduct of experiments, then, concerns the way human and animal experiments should be conducted, and whether they are permissible for the purpose at all. Grüne-Yanoff goes on to explain how students are often taught only the rules of ethical conduct and scientific honesty. He also argues that students should instead learn to understand, critically reflect, and apply these rules. He proposes to help students to elicit intuitive judgments about concrete cases. The role of the philosopher would then be to analyze these judgments in terms of their specific values (e.g., honesty, openness, transparency, confidentiality) (Grüne-Yanoff, 2014). We have developed such an assignment in our undergraduate program which could illustrate the type of training we aim for. For this assignment, students are asked to collectively rank several different life forms in terms of their moral status (e.g., human cell line, organoid, 3-day-old embryo, 9-week-old fetus, test rat, own domestic rat, a brain-dead person, etc.). Then, they are asked to justify their ranking, discuss what possible reasons for disagreement were, and describe which characteristics of the life form they considered (e.g., potential of having human characteristics, self-awareness, personal relationship, etc.). As teachers, we simultaneously analyze the types of argument they put forward and help students with describing in more general terms how they came to their ranking. With this assignment, we ask students to analyze their own moral judgment (and we supplement that analysis when necessary) to teach them methods of analysis of moral judgment.

Teachers can have a similar role when discussing *consequences* of research. Philosophy can facilitate reflection on what bad consequences can be avoided and what good can be done (Grüne-Yanoff, 2014). Grüne-Yanoff, here, proposes to use historic examples

[4] As a side note, Grüne-Yanoff (2014) puts scientific honesty under non-epistemic values while some aspects of honesty help toward the attainment of epistemic goals as well and might as such be viewed as an epistemic value.

and thought experiments to illustrate what such consequences might be. We add that it is important to get students into the habit of asking these questions, not only when it is specifically asked of them during a case study assignment for which reflection on the consequences of research is the explicit objective. These ways of thinking and acting should also be integrated into biomedical courses. We need researchers to start attending to the non-epistemic aspects of their research on their own, routinely. This means that teachers also need to provide feedback about these aspects during assignments for which reflection on the consequences of research is not explicitly requested. We cannot expect students to automatically transfer their case study-based skills learned during ethics education to their own scientific practice. We need to support this transfer.

An example of how non-epistemic considerations can impact research decisions can be found in the use of omics data. These data can be so unique that people can be reidentified from their data even when other identifiable data are not disclosed. To conduct research on these data ethically responsibly, the privacy of the subjects must be protected. This needs to be weighed against the epistemic value of reproducibility, which is enhanced by openly sharing research data. Ethical reflection, but also legislation often comes too little too late, after the impacts have already been realized. This is especially disconcerting for biomedical data science, because it is a quickly evolving research area with profound impact on healthcare and society. Thus, timely reflection on the effects of non-epistemic considerations is especially important in this field and should therefore be taught as well. So, when discussing data management and the advantages of Open Science during biomedical data science education, we should also discuss the non-epistemic aspects and their consequences.

Another non-epistemic aspect discussed by Grüne-Yanoff (2014) is the connection between scientific research and policy making. He mentions the discrepancy between judging evidence for epistemic or for policy purposes. For example, in setting the type I (false positive) and type II (false negative) error margins. From an epistemic perspective we often prefer a minimal type I error, while from many policy perspectives a small type II error may be preferred. Grüne-Yanoff then concludes that philosophy can help students to analyze such problems through the tools of decision theory (Grüne-Yanoff, 2014). He does not elaborate on how this can take shape in a (philosophy course in a) science curriculum. For an undergraduate program, which we discuss, it is probably too complicated and time-consuming to introduce decision theory (tools). However, the realization that what is decisive when making a choice in a research process depends on the perspective from which the consequences are assessed is important. From that point of view, not only political perspectives are important for a biomedical scientist. In addition to epistemic and political perspectives, biomedical scientists should, for example, also consider clinical, legal, participant, data management, societal, organizational, pragmatic, sustainability, and financial perspectives. Further, they need to become aware of their own perspectives and varying roles as well. In addition to their role as researchers, scientists fulfill other roles as well. For example, they have an expert-role in the political and societal domain (Grüne-Yanoff, 2014) and they might have an advisory role in NGOs or companies. This means that in addition to external perspectives, a scientist should consider their own varying perspectives as well during decision-making. Therefore, we suggest that undergraduate students learn to have an eye for these perspectives first. Again, we do not suggest that undergraduate students should oversee the consequences of their decisions from all these perspectives. Rather, they should learn to *inquire about* these consequences with their teachers and supervisors first. Once they have become aware of these perspectives as undergraduates, they can learn to integrate these perspectives in their own decision-making during their master's or PhD. This requires an introduction to the perspectives early in the curriculum and consolidation throughout the rest of the program. The latter could be implemented by adding a standard question about perspectives to assignments concerning decision-making in research.

Being aware of how non-epistemic considerations can affect scientific research, again, has to do with understanding the nature of science (fourth dimension of learning). As can be appreciated from the discussion above, the proposed training can provide an opportunity to discuss the human-constructed, subjective, and socio-cultural embedded nature of science and scientific knowledge. However, it is good to note this should be done explicitly and reflectively (Abd-El-Khalick & Lederman, 2000). We should not expect students to learn about the nature of science automatically by participating in research practices (Abd-El-Khalick & Lederman, 2000). We need to take the opportunities we present here to explicitly discuss how students view the nature of science and how they reflect on their own research practice.

To summarize, the learning goals suggest students should learn to act on their ethical and societal responsibilities and the effects of their decisions on people. We add that students should learn to reflect on the effects of ethics and society on scientific decision-making and that students need to learn to balance their epistemic, ethical, and societal responsibilities. We suggest teaching students how to identify underlying values of epistemic and non-epistemic considerations. And teaching them to reflect on effects of non-epistemic aspects of decision-making by making them aware of their own roles and other perspectives on scientific research.

4.5 Key Aspect 5: Report Justification of Decisions and Results Appropriately

The fifth key aspect of conducting research that we identified in the learning goals pertains to reporting scientific research to the scientific community. In the bioinformatics learning goals, we can read that "[An undergraduate is learning] both to recognize the value of clear communication, & about the role of communication in sharing & publishing research" (Tractenberg et al., 2019, p. 12). As a side note, it is good to mention that "communication of scientific research" in these learning goals is focused on communication between scientists and not necessarily with the general public. The latter is often regarded as the field of science communication. Skills in science communication are also very valuable for (future) biomedical scientists. However, this is beyond the scope of this article. The learning goals discussed here are focused on reporting of scientific research to peers, for example, in scientific journals. In this respect, they suggest that an undergraduate is learning "that transparency in all communication represents ethical practice, even when the desired results have not been achieved" (Tractenberg et al., 2019, p. 12). Further, students need to learn "why *p*-value-driven conclusions, & the lack of false discovery rate controls, are not conducive to reproducible work" (Ibid.). To summarize these learning goals, students need to learn what constitutes a complete report of research. In addition, the learning goals suggest that an undergraduate's "conclusions are generally aligned with given results" and that undergraduates "with guidance, can draw conclusions in own work that are coherent with the research hypothesis/hypotheses" (Ibid.). Further, an undergraduate is "learning that 'full' contextualization of conclusions requires consideration of limitations deriving from methods & their applications, & their effects on results & conclusions" (Ibid.). In other words, students need to learn to align and contextualize their conclusions to provide a congruent report. Thus, they need to learn that reports of research should include all aspects relevant to assess the conclusions drawn and that this includes research (design) decisions and how they affect results and conclusions. The last key aspect of conducting research we can now identify is that students should learn to "report justification of decisions and results appropriately."

4.5.1 Train Presentation of Research Reports as Arguments

As Tractenberg (2018) already clarifies, the key goal for teaching a student to draw and contextualize conclusions is to teach them to support an argument with or from data by *reasoning*. Beginners often completely rely on *p*-values, instead of using reasoning, to support an argument (Tractenberg, 2018). So, students should progress from relying solely on *p*-values to using reasoning for drawing conclusions.

We add that this progress should also be sought in *reporting* conclusions. Reporting of results should reflect the reasoning that led to the conclusions. To bring about that progress, we should focus on the underlying assumptions of students about conclusions in scientific research. In our experience, many students believe that scientific research provides evidence for a conclusion. And, indeed, that finding a *p*-value below the significance level is sufficient evidence to conclude that the hypothesis is correct and relevant. There are two assumptions that should be corrected: (1) that the evidence speaks for itself, and (2) that one can draw conclusions about the truth of a hypothesis (from a single well-performed study).

Students should first be aware of whether they hold these beliefs. Then, insights from philosophy of science can help to understand that the value of the evidence (e.g., a *p*-value or an effect size) is reliant on many factors, such as the research methods, theory-ladenness, and social and cultural factors (as addressed by the fourth dimension of learning). Consequently, students can more thoroughly understand why it is important to view reporting scientific results as constructing an argument. And that this argument includes the *p*-value, but also the decisions that were made and every result that was obtained to arrive at that *p*-value. And that it includes other results of research done on this topic. So, viewing research as constructing an argument instead of letting the evidence speak for itself could help students to both be congruent and complete in communicating their research results. And to know more intuitively what constitutes a complete and congruent research report, without having to learn a list of specific requirements, such as detailed in the learning goals.

4.5.2 Train Reflection on the Epistemic Status of Hypotheses in Research Reports

In "Sect. 4.3," we already suggested that appreciating the value of taking a step and providing evidence that supports or does not support a hypothesis *to a certain extent* can help students with explaining how decisions affect results. Here, we add that it might decrease their tendency to use reasoning fallacies to give the impression that their results are conclusive. So, reflecting on the epistemic status of hypotheses and conclusions in research reports could help students to avoid reasoning fallacies. Therefore, training in taking apart and constructing arguments could prove to be very useful for appropriate reporting of scientific research.

One of the reasons viewing research efforts as constructing arguments is valuable for research on data of increasing size and complexity is because there is an increased risk of confusing hypothesis testing and hypothesis generation efforts. Big data can be used to mine for new hypotheses, but this is a challenging exercise (Tatonetti, 2019). For example, the large sample size can cause the study to be overpowered to detect extremely small effect sizes with statistical significance. In addition, there is a larger risk of confounding in large, observational datasets. Therefore, validation of these findings from machine learning is essential. However, validation is not always performed on truly independent data, which leads to inflated results (Tatonetti, 2019). Hypothesis generation and hypothesis testing efforts have different argumentative values that should complement each other. Recognizing hypothesis generation as such will put the argumentative value of its results into perspective and attest to the importance of subsequent hypothesis testing efforts.

Reflection on the epistemic status of hypotheses in research reports can be integrated in several biomedical data science learning activities. For example, during journal clubs in which bioinformatics papers are discussed, we should include questions about the argument that is built in these papers. Journal clubs often address the findings and figures of a paper and are used to discuss the structure of research articles. To prepare for these journal clubs, students are given a reading instruction with guiding questions to focus their preparation. During the journal club these questions are then discussed with peers and a teacher. To foster the idea that reporting research is constructing an argument, teachers could add questions about the persuasiveness of the paper. For example, by asking students what, in their opinion, contributes to the persuasiveness and what does not. Further, by discussing what, in subsequent research, needs to be done to test the explanatory or predictive power of the theory under development further. Or, by asking which choices were made by the researchers and how these choices affected theory building. These exercises can foster an understanding of the tentative and human-constructed nature of scientific knowledge. In addition, it provides room to discuss that data resulting from observations or experimentation require interpretation to build theories. This difference between observations and inferences is another aspect of understanding the nature of science that is promoted by nature of science scholars (Lederman et al., 2002).

Another example of a relevant learning activity is the thesis at the end of a degree program. Usually, students conduct a small empirical study under supervision of a researcher, and they write a report in the form of a research article. Since students are actually engaged in scientific practice during their research internship, there are many opportunities for supervisors to address the tentative and human-constructed nature of scientific knowledge. However, it is good to note again that this should be done explicitly and reflectively (Abd-El-Khalick & Lederman, 2000). Here, we recognize the importance of the fourth dimension of learning again. We propose, for example, that supervisors let their students reflect on the objectives of reporting research in scientific journals. This could provide an opportunity to uncover and address students' assumptions (e.g., that data speak for themselves). As nature of science scholars have expressed, students and teachers with better understanding of the nature of science will argue with more evidence (McDonald, 2010). However, teacher guidance is necessary to help students apply their understanding of the nature of science and to appreciate its relevance to effective argumentation (McDonald, 2010). Further, instructing students to disassemble their own argument in their thesis could help them identify their own assumptions and reasoning fallacies.

To summarize, the learning goals focus on the importance of congruency and completeness and using reasoning for reporting of scientific research. This implies the philosophical skill of argumentation. To help students achieve these learning goals, we propose to help students develop the following view on scientific research: scientific research is constructing an *argument* to demonstrate *to what extent* some evidence supports a conclusion. Focusing on the argumentative aspects of reporting scientific research and training in

Table 2 Philosophical training for biomedical sciences undergraduates

Key aspects of conducting scientific research	Philosophical training	Associated aspects of the nature of science (NOS)
1. Understand data, research methods, and statistical methods, and use them	This aspect of conducting research benefits from attending to the other aspects by focusing on ways of thinking and acting (higher dimensions of learning) that transcend practical actions	Empirical NOS Human-constructed NOS Difference between observations and inferences
2. Make and justify decisions	Discuss why relevance, assumptions, uncertainties, reproducibility, and rigor are relevant considerations in scientific decision-making Make considerations explicit in scientific decision-making Identify the principles, beliefs, and values that guide decision-making by taking the process apart Reflect on the role of probability and uncertainty in science Reflect on the assumptions behind statistical concepts	Human-constructed NOS Socio-cultural embeddedness of science Intersubjective, collaborative NOS
3. Explain how decisions affect results	Discuss how epistemic aspects of decision-making affect scientific results Discuss how scientists are part of a disciplinary paradigm and how that affects the epistemic status of current scientific knowledge Discuss how the value of evidence is reliant upon research methods, theory-ladenness, and social and cultural factors Discuss how scientific research contributes to theory building Discuss how scientific research provides evidence that supports a theory to a certain extent (instead of proving or disproving theories)	Tentative NOS Human-constructed NOS Focus on explanatory and predictive value of scientific theories (rather than truth value)

Table 2 (continued)

Key aspects of conducting scientific research	Philosophical training	Associated aspects of the nature of science (NOS)
4. Balance epistemic, ethical, and societal responsibilities	Reflect on the responsibilities of scientific researchers toward society Balance non-epistemic with epistemic considerations, by understanding their underlying values Discuss how non-epistemic aspects affect science Reflect on the effects of ethical conduct, the societal implications of research, and the scientist's expert role on scientific research itself Analyze moral judgments Inquire about consequences of research for various stakeholders by perspective-taking	Human-constructed NOS Subjective NOS Socio-cultural embeddedness of science
5. Report justification of decisions and results appropriately	Discuss how reporting of results is constructing an argument (instead of letting the evidence speak for itself) Reflect on the epistemic status of hypotheses in research reports Discuss the role of making and justifying decisions, researcher interpretation, and persuasiveness in research articles	Tentative NOS Empirical NOS Human-constructed NOS Difference between observations and inferences

taking apart and constructing arguments could help students to report results appropriately by using reasoning and avoiding reasoning fallacies.

A summary of all our proposed training in philosophy of science for learning to conduct scientific research through higher dimensions of learning can be found in Table 2.

5 Other Benefits of Integration of Philosophy of Science in Biomedical Data Science Education

Thus far, we have focused on the benefits of integration of philosophy of science in biomedical data science education for learning biomedical data science. Another benefit of this integration, which we discussed between the lines, is its value for learning philosophy of science. As is also discussed by Boniolo and Campaner (2020), philosophers of science should teach biomedical students what they need to better understand the foundations of their theoretical knowledge and practice. Grüne-Yanoff (2014) describes a multidisciplinary course to introduce students from different disciplines to philosophy of science. As is acknowledged, it is a constraint that students come from highly diverse backgrounds (Grüne-Yanoff, 2014). Since the educational content needs to be tailored to all students, many examples will not be (directly) relevant to most students. In our experience, the philosophical concepts that help to understand the foundations of scientific theory and practice can lose their practical value for students when they are taught separate from particular biomedical content and research techniques. However, when the concepts that are applicable to some disciplinary content are taught when that content is also covered from a disciplinary perspective, the chances of them taking root are significantly increased. In addition, it can increase student interest for theoretical concepts. As we have argued above, biomedical data science is a fitting subject for integration of philosophy of science training. The core competencies as defined by Wilson Sayres et al. (2018) provide an excellent overview of important datatypes and research and analysis methods in biomedical data science. These can be used to integrate the educational approaches discussed here into existing curricula. When the philosophical concepts and skills are taught together with disciplinary content it also facilitates their actual application in research practice. On a more practical note, it often proves difficult to create space that is explicitly dedicated to philosophy of science in a packed curriculum. An integration with disciplinary content could, therefore, also be a pragmatic solution. So, integration could be practical, increase comprehension of and motivation for learning philosophical concepts, and facilitate application of these concepts in research practice.

An additional benefit of the discussed educational strategies is their value for interdisciplinary research efforts. Integration of interdisciplinarity is also one of the skills that is described by Tractenberg et al. (2019). However, as becomes apparent in their explication of the learning goals they use a narrow view on interdisciplinarity. The domains that are mentioned are biomedicine, statistics, and engineering. As has been argued by others, the complex problems of our time require broader interdisciplinary efforts (e.g., Mazzocchi, 2019). Biomedical scientists should, for example, also collaborate with sociologists, psychologists, jurists, and communication experts. The differences between biomedicine and these domains concerning methodology and the underlying principles and values are even starker. Therefore, collaborations between these domains require greater interdisciplinary skills.

Table 3 Four initial design principles

Design principle	Reasons
1. Integrate training in philosophy of science with biomedical data science training	Making philosophy of science meaningful Fostering a habit of attending to philosophical aspects while conducting research
2. Attend simultaneously to all four dimensions of learning to conduct scientific research. While attending to decisions in research design and implementation, also explicitly reflect with students on their views of the nature of science	Understanding rationales instead of learning disciplinary conventions -Learning about aspects of the nature of science in a meaningful context
3. Focus on aspects of the *process* of scientific decision-making and explicitly address students' possible tendency to focus on the outcomes of scientific decision-making	Learning that multiple decisions can be right instead of seeking for "correct" decisions Learning that justification of decisions and consideration of their consequences is important
4. Teach students to ask questions about epistemic and non-epistemic considerations in scientific research	Learning that scientific research practice is defined by many choices between reasonable alternatives Developing awareness of these choices instead of learning predefined epistemic and ethical rules

Philosophy of science can strengthen these skills. For example, being able to abstract the considerations for decisions in one's own domain and understanding the principles of argumentation are valuable skills for interdisciplinary projects. They can help students to justify their proposed research methodology for collaborative projects in a manner that is understandable for practitioners of other domains (Grüne-Yanoff, 2014; Laplane et al., 2019). Furthermore, with these skills students are also better equipped to understand the underlying principles of research and analysis methods of these other domains. Thus, philosophy of science provides a common language for interdisciplinary conversations, which is important given the complex problems of our time.

In addition, other disciplines can provide tools for biomedical scientists to improve their own practice in a comparable manner as we argued for philosophy of science tools here. Above, we already touched upon the fields of ethics, science communication, and cognitive science. Philosophical tools can provide insight into ethical decision-making, communication of scientific research, and the ways in which we convey and perceive information through language (e.g., statistical terms). Naturally, the fields of ethics, science communication, and cognitive science have also developed their own tools and insights on these topics. Integration of these tools and insights into biomedical data science education could provide similar usefulness as we have argued for integration of philosophy of science.

6 Educational Design Principles

Throughout our article we have provided recommendations for how to help students achieve a way of thinking and acting that can improve their scientific practice, especially when working with increasingly large and complex datasets. We summarize these recommendations in four educational design principles (Table 3). The first design principle is an overarching one that has been the core of our argument: Integrate training in philosophy of science with biomedical data science training. Philosophical exercises are most meaningful when they are done at times when the content they address is also addressed from a disciplinary perspective. In addition, it fosters a habit of attending to philosophical aspects while conducting research because philosophical thinking is already presented as integral to conducting research during formal education.

The second design principle derives from our identification of four dimensions of learning to conduct scientific research. When teaching data science to undergraduates, we tend to focus on introducing students to a wide range of biomedical data science tools and workflows. Like with other disciplinary content, students are taught disciplinary conventions (Grüne-Yanoff, 2014). Even during internships, undergraduates are often used as "extra hands" and gain little experience in making well-considered choices themselves (Li & Luo, 2020). Furthermore, little attention is paid to reflection on the nature of science. However, through an understanding of the nature of science, researchers can understand rationales behind choices in research design and implementation, make justified choices, and apply techniques in a meaningful manner. Vice versa, aspects of the nature of science only make sense when applied to specific choices and applications. Therefore, our second design principle is: Attend simultaneously to all four dimensions of learning to conduct scientific research. While attending to decisions in research design and implementation, we should explicitly reflect with students on their views of the nature of science.

Another point of attention for biomedical data science education is that students might have the wrong impression that teachers seek "correct" decisions or the "right" answers to their questions when teaching about scientific research. Students need to realize that multiple decisions can be right, and that justification of decisions and consideration of their consequences is more important. This requires explicit consideration. Therefore, our third design principle is: Focus on aspects of the *process* of scientific decision-making and explicitly address students' possible tendency to focus on the *outcomes* of scientific decision-making.

Lastly, when discussing considerations in scientific research, we should shift our focus from teaching predefined epistemic and ethical rules and guidelines to teaching students to ask relevant questions during decision-making. By learning rules and guidelines, students might get the wrong impression that appropriate conduct of research is unequivocal. Actual scientific research practice, however, is defined by many choices between reasonable alternatives. This applies to both epistemic and non-epistemic aspects of scientific research. Appropriate conduct of research is, among others, characterized by an awareness of these choices and reflection on important considerations before making choices. Undergraduate students might not be able to oversee all important considerations nor the implications of different decisions. They can, however, be stimulated to take responsibility for scientific decision-making when they know which questions to ask. Therefore, our fourth design principle is: Teach students to ask questions about epistemic and non-epistemic considerations in scientific research.

7 Conclusions and Discussion

We argued that integration of training in philosophy of science in biomedical data science education can increase the achievability of biomedical data science learning goals. We advocated for training in the explication of considerations for scientific decisions, in reflection on the principles of probability and statistics, and in the inclusion of non-epistemic considerations for decisions. These skills could help students to make and justify decisions and understand their implications. Then, we made a case for a focus on the argumentative aspects of reporting in science and training in taking apart and constructing arguments to help students to report results in a congruent and complete manner. Together, this training on higher dimensions of learning to conduct scientific research contributes to achieving competency in the understanding and appropriate use of biomedical data and scientific methods. With the discussed philosophical training it will be easier for students to gain knowledge about biomedical data science methods and to appropriately use statistical and computational tools. Furthermore, by tapping into higher dimensions of learning, students learn transferable skills that can be used to evaluate unencountered types of inquiry as well.

When one studies the general principles, beliefs, and values that guide scientists in making decisions, it will become apparent that there are inherent uncertainties and that scientists do operate in a paradigm. Reflection on these principles, beliefs, and values by students themselves can help to let these insights in the nature of science take root. Especially when they are taught hand in hand with particular biomedical content and research techniques. Consequently, the chances of students using these insights to guide their scientific practice will increase.

The proposed focus of educational strategies on higher dimensions of understanding of rigorous and reproducible research is particularly valuable for the field of biomedical data science. The increasing size and complexity of biomedical data require an increasing number of different data analysis methods and development of new methods. Therefore, students should be equipped to learn to correctly use and develop new methods after their studies. This requires being able to make (considerations for) decisions in scientific research explicit and being able to break down different methods. This can be trained through reflection on the nature of science, its purpose, scope, and methods, and on justifications of the scientific ways of proceeding. Another aspect of data science that benefits from a conceptual understanding of science and its argumentative nature is the increased risk of confusing hypothesis testing and hypothesis generation efforts. It is important to recognize the argumentative value of results of hypothesis generation efforts, to avoid common pitfalls in their use.

Our identified dimensions of learning increase from practical aspects to a more conceptual understanding. In that sense, they resemble the dimensions of scientific literacy proposed by Bybee (1997). The dimensions of scientific literacy he discerns are nominal, functional, conceptual and procedural, and multidimensional. In short, nominal literates can recognize scientific terms, functional literates can use simple scientific vocabulary, conceptual and procedural literates understand scientific concepts and the scientific procedures for developing new explanations and inventions, and multidimensional literates understand conceptual structures and the philosophical, historical, and social dimensions of science (Bybee, 1997). However, scientific literacy as a concept and Bybee's definition of it are focused on *general* education. This means that designing for scientific literacy should begin "by asking what it is a student ought to know, value, and do as a citizen" (Bybee, 1997, p. 73). Therefore, the dimensions of scientific literacy are focused on *consumers* of science rather than *producers* of science, which is what we focus on in this article. Nonetheless, it could be further explored how the parallel between the dimensions of learning to conduct scientific research and the dimensions of scientific literacy can inform educational approaches for university degree programs.

Throughout our discussion of educational approaches, we have expressed our concern that undergraduate students (and other biomedical scientists) are under the impression that science is about truth and good science leads to unambiguous results. That students and science teachers hold these positivistic, idealistic views of science is also supported by others (Abd-El-Khalick & Lederman, 2000; Howitt & Wilson, 2018). However, these studies of views of the nature of science have mainly been conducted among primary and high school students, and their science teachers (Deng et al., 2011). In addition, many of these studies characterized students' professed views, when the nature of science was explicitly addressed in the classroom or in an educational research setting (Deng et al., 2011). However, the views students say they hold when asked directly might not be the same as the views they *enact* in their own scientific practice or in interpreting the products and procedures of science in everyday practice. Therefore, to enable educators to support conceptual change in students, we first need to better characterize how university students think about the nature of science and what views they enact in their scientific practice.

Another point of attention with our proposed educational approach is that we rely on biomedical teachers without a philosophy of science background to integrate these ways of thinking and acting in their learning activities. We believe this integration in courses focused on biomedical (data) science content is essential. However, it requires training of biomedical teachers as well. Another solution to this problem might be found in direct collaborations between biomedical and philosophy of science teachers. In such a co-teaching approach, biomedical scientists and philosophers of science can combine their expertise in front of the classroom. This could create a lesson which attends to the higher dimensions of learning with direct application to biomedical content. In addition, it provides an opportunity for biomedical scientists and philosophers of science to learn from each other. It could meet the need of biomedical scientists to gain deeper understanding of the nature of their scientific practice and the need of philosophers of science to learn what scientists really need from philosophy of science. Whether and how co-teaching can accommodate integration of philosophy of science in biomedical data science education should be further explored.

Similarly, we rely on biomedical scientists in the labs where students gain research experience. Students might be trained like we propose in their formal curriculum, but it might be difficult to let these habits take root in their scientific practice. Due to (perceived) hierarchy, students might conform to the practices in the lab they are working in. They might not feel confident enough, for example, to address the epistemic, ethical, and societal responsibilities they perceive as conflicting in their (supervisor's) research approach. These research experiences might be more influential than what they learn in their formal curriculum. Again, training in philosophy of science could also prove to be useful in continued professional education of scientists, for example, during seminars or conferences. These researchers play a crucial role in how students form their identity as scientists.

Lastly, it is good to note that with our focus on the second to fifth aspect of conducting research, we do not mean to imply that we should not train students in the use of biomedical data science methods and tools. Rather, what we propose is a shift of focus. We need to choose what to spend our time in formal education of biomedical scientists on and should focus on the aspects that make for good scientists. We believe that a focus on understanding of the rationale for choices in research design and implementation and understanding the nature of science contributes more to that goal than emphasis on expanding a student's biomedical data science toolkit. Even so, training in philosophy of science cannot prevent questionable research practices on its own. Rather, it is one of the paths to reform scientific practice. Other reforms are necessary. For example, a reform of the reward culture of science, revision of the statistical inferential framework, promoting replications, publishing negative findings, and Open Science initiatives. Here, we presented one way to reform scientific practice through formal education.

In conclusion, we proposed to bring the objectives of biomedical data scientists and those of philosophers of science for life science education together. We have shown that these objectives are aligned and could be mutually reinforcing when taught together. The integrated knowledge, skills, and abilities can better prepare students for understanding and researching increasingly complex biomedical data with the purpose of creating knowledge about human life, health, and disease.

Acknowledgements We thank Rens van de Schoot for useful discussions about this article.

Author Contribution Annelies Pieterman-Bos had the idea for the article, performed the literature search and analysis, and drafted the work. Marc H. W. van Mil guided the elaboration of the idea and critically revised the work.

Funding This study was funded by Utrecht University.

Data Availability Not applicable.

Code Availability Not applicable.

Declarations

Conflict of Interest The authors declare that they have no conflict of interest.

References

Abd-El-Khalick, F. (2012). Nature of science in science education: Toward a coherent framework for synergistic research and development. In B. J. Fraser, K. G. Tobin, & C. J. McRobbie (Eds.), *Second international handbook of science education* (pp. 1041–1060). Springer. https://doi.org/10.1007/978-1-4020-9041-7_69

Abd-El-Khalick, F., & Lederman, N. G. (2000). Improving science teachers' conceptions of nature of science: A critical review of the literature. *International Journal of Science Education, 22*(7), 665–701. https://doi.org/10.1080/09500690050044044

Altman, R. B., & Levitt, M. (2018). What is biomedical data science and do we need an annual review of it? *Annual Review of Biomedical Data Science, 1*(1), i–iii. https://doi.org/10.1146/annurev-bd-01-041718-100001

Attwood, T. K., Blackford, S., Brazas, M. D., Davies, A., & Schneider, M. V. (2019). A global perspective on evolving bioinformatics and data science training needs. *Briefings in Bioinformatics, 20*(2), 398–404. https://doi.org/10.1093/bib/bbx100

Attwood, T. K., Bongcam-Rudloff, E., Brazas, M. E., Corpas, M., Gaudet, P., Lewitter, F., Mulder, N., Palagi, P. M., Schneider, M. V., van Gelder, C. W. G., Attwood, T. K., Banag, C., Blackford, S., Blatter, M. C., Bongcam-Rudloff, E., Brazas, M. D., Brooksbank, C., Budd, A., Charleston, M., & Zanzoni, A. (2015). GOBLET: The Global Organisation for Bioinformatics Learning Education and Training. *Plos Computational Biology, 11*(4), 1–10. https://doi.org/10.1371/journal.pcbi.1004143

Benchimol, E. I., Smeeth, L., Guttmann, A., Harron, K., Moher, D., Petersen, I., Sorensen, H. T., von Elm, E., Langan, S. M., Working, R. E. C. O. R. D., & Committee. (2015). The reporting of studies conducted using observational routinely-collected health data (RECORD) statement. *PLOS Medicine, 12*(10), e1001885. https://doi.org/10.1371/journal.pmed.1001885

Boniolo, G., & Campaner, R. (2020). Life sciences for philosophers and philosophy for life scientists: What should we teach? *Biological Theory, 15*(1), 1–11. https://doi.org/10.1007/s13752-019-00333-7

Braillard, P.-A. (2013). Systems biology and education. In K. Kampourakis (Ed.), *The philosophy of biology: A companion for educators* (pp. 549–575). Springer. https://doi.org/10.1007/978-94-007-6537-5_24

Bybee, R. W. (1997). *Achieving scientific literacy: From purposes to practices*. Heinemann

Corpas, M., Jimenez, R. C., Bongcam-Rudloff, E., Budd, A., Brazas, M. D., Fernandes, P. L., Gaeta, B., Van Gelder, C., Korpelainen, E., Lewitter, F., McGrath, A., MacLean, D., Palagi, P. M., Rother, K., Taylor, J., Via, A., Watson, M., Schneider, M. V., & Attwood, T. K. (2015). The GOBLET training portal: A global repository of bioinformatics training materials, courses and trainers. *Bioinformatics, 31*(1), 140–142. https://doi.org/10.1093/bioinformatics/btu601

Deng, F., Chen, D. T., Tsai, C. C., & Chai, C. S. (2011). Students' views of the nature of science: A critical review of research. *Science Education, 95*, 961–999. https://doi.org/10.1002/sce.20460

Dill-McFarland, K. A., Konig, S. G., Mazel, F., Oliver, D. C., McEwen, L. M., Hong, K. Y., & Hallam, S. J. (2021). An integrated, modular approach to data science education in microbiology. *PLoS Computational Biology, 17*(2), e1008661, . https://doi.org/10.1371/journal.pcbi.1008661

Dinsdale, E., Elgin, S. C. R., Grandgenett, N., Morgan, W., Rosenwald, A., Tapprich, W., Triplett, E. W., & Pauley, M. A. (2015). NIBLSE: A network for integrating bioinformatics into life sciences education. *CBE Life Sciences Education, 14*(4), 1–4. https://doi.org/10.1187/cbe.15-06-0123

Elliott, J. H., Grimshaw, J., Altman, R., Bero, L., Goodman, S. N., Henry, D., Macleod, M., Tovey, D., Tugwell, P., White, H., & Sim, I. (2015). Make sense of health data. *Nature, 527*(7576), 31–32. https://doi.org/10.1038/527031a

Ernst, Z. (2009). Philosophical issues arising from genomics. In M. Ruse (Ed.), *The Oxford handbook of philosophy of biology* (pp. 310–348). Oxford University Press. https://doi.org/10.1093/oxfordhb/9780195182057.003.0014

Gerrits, E. M., Bredenoord, A. L., & van Mil, M. H. W. (2021). Educating for Responsible Research Practice in Biomedical Sciences: Towards Learning Goals. *Science & Education*. https://doi.org/10.1007/s11191-021-00295-y

Grüne-Yanoff, T. (2014). Teaching philosophy of science to scientists: Why, what and how. *European Journal for Philosophy of Science, 4*(1), 115–134. https://doi.org/10.1007/s13194-013-0078-x

Hicks, S. C., & Irizarry, R. A. (2018). A guide to teaching data science. *The American Statistician, 72*(4), 382–391. https://doi.org/10.1080/00031305.2017.1356747

Horton, N. J. (2015). Challenges and opportunities for statistics and statistical education: Looking back, looking forward. *American Statistician, 69*(2), 138–145. https://doi.org/10.1080/00031305.2015.1032435

Howitt, S. M., & Wilson, A. N. (2018). Reflecting on the use and abuse of scientific data facilitates students' ethical and epistemological development. *Science Education, 102*(3), 571–592. https://doi.org/10.1002/sce.21333

Kleinschmit, A. J., Brink, B., Roof, S., Goller, C. C., & Robertson, S. (2019). S*equence similarity: An inquiry based and "under the hood" approach for incorporating molecular sequence alignment in introductory undergraduate biology courses* [Teaching resource]. CourseSource. https://doi.org/10.24918/cs.2019.5

Laplane, L., Mantovani, P., Adolphs, R., Chang, H., Mantovani, A., McFall-Ngai, M., Rovelli, C., Sober, E., & Pradeu, T. (2019). Why science needs philosophy. *Proceedings of the National Academy of Sciences of the United States of America, 116*(10), 3948–3952. https://doi.org/10.1073/pnas.1900357116

Lederman, N. G. (2007). Nature of science: Past, present, and future. In S. K. Abell & N. G. Lederman (Eds.), *Handbook of research on science education* (pp. 831–879). Lawrence Erlbaum.

Lederman, N. G., Abd-El-Khalick, F., Bell, R. L., & Schwartz, R. S. (2002). Views of nature of science questionnaire: Toward valid and meaningful assessment of learners' conceptions of nature of science. *Journal of Research in Science Teaching, 39*(6), 497–521. https://doi.org/10.1002/tea.10034

Li, J., & Luo, L. (2020). Nurturing undergraduate researchers in biomedical sciences. *Cell, 182*(1), 1–4. https://doi.org/10.1016/j.cell.2020.05.008

Lindley, D. V. (2000). The philosophy of statistics. *Journal of the Royal Statistical Society. Series D (The Statistician)*, *49*(3), 293–337. https://www.jstor.org/stable/2681060

Madlung, A. (2018). Assessing an effective undergraduate module teaching applied bioinformatics to biology students. *PLoS Computational Biology, 14*(1), 1–15. https://doi.org/10.1371/journal.pcbi.1005872

Mayer, M. (2018). Research integrity and the law that never was. *BMJ Evidence-Based Medicine, 23*(6), 218–224. https://doi.org/10.1136/bmjebm-2018-110993

Mazzocchi, F. (2019). Scientific research across and beyond disciplines. *EMBO Reports, 20*, e47682. https://doi.org/10.15252/embr.201947682

McComas, W. F. (2020). Principal elements of nature of science: Informing science teaching while dispelling the myths. In W. F. McComas (Ed.), *Nature of science in science instruction: Rationales and strategies* (pp. 35–65). Springer. https://doi.org/10.1007/978-3-030-57239-6_3

McDonald, C. V. (2010). The influence of explicit nature of science and argumentation instruction on preservice primary teachers' views of nature of science. *Journal of Research in Science Teaching, 47*(9), 1137–1164. https://doi.org/10.1002/tea.20377

Montgomery, E. B. (2019). Reproducibility in biomedical research: Epistemological and statistical problems. *Elsevier Science*. https://doi.org/10.1016/C2018-0-02296-3

Nie, X., Zhang, Y., Wu, Z., Jia, L., Wang, X., Langan, S. M., Benchimol, E. I., & Peng, X. (2018). Evaluation of reporting quality for observational studies using routinely collected health data in pharmacovigilance. *Expert Opinion on Drug Safety, 17*(7), 661–668. https://doi.org/10.1080/14740338.2018.1484106

Olimpo, J. T., Pevey, R. S., & McCabe, T. M. (2018). Incorporating an interactive statistics workshop into an introductory biology course-based undergraduate research experience (CURE) enhances students' statistical reasoning and quantitative literacy skills. *Journal of Microbiology & Biology Education, 19*, 1. https://doi.org/10.1128/jmbe.v19i1.1450

Rosenwald, A. G., Pauley, M. A., Welch, L., Elgin, S. C. R., Wright, R., & Blum, J. (2016). The CourseSource bioinformatics learning framework. *CBE Life Sciences Education, 15*(1), 1–2. https://doi.org/10.1187/cbe.15-10-0217

Schwartz, R., & Lederman, N. (2008). What scientists say: Scientists' views of nature of science and relation to science context. *International Journal of Science Education, 30*(6), 727–771. https://doi.org/10.1080/09500690701225801

Tatonetti, N. P. (2019). Translational medicine in the age of big data. *Briefings in Bioinformatics, 20*(2), 457–462. https://doi.org/10.1093/bib/bbx116

Tractenberg, R. E. (2017). How the mastery rubric for statistical literacy can generate actionable evidence about statistical and quantitative learning outcomes. *Education Sciences, 7*(1), 3. https://doi.org/10.3390/educsci7010003

Tractenberg, R. E. (2018). *Training with the mastery rubric for statistical literacy to promote rigor and reproducibility across scientific disciplines: Making the journal club educational* [Proceedings]. Joint Statistical Meetings, Vancouver, BC, Canada.

Tractenberg, R. E., Lindvall, J. M., Attwood, T. K., & Via, A. (2019). The mastery rubric for bioinformatics: A tool to support design and evaluation of career-spanning education and training. *PLoS One, 14*(11), 1–29. https://doi.org/10.1371/journal.pone.0225256

Valcu, M., & Valcu, C. M. (2011). Data transformation practices in biomedical sciences. *Nature Methods, 8*(2), 104–105. https://doi.org/10.1038/nmeth0211-104

Vaught, M., Jordan, D. C., & Bastian, H. (2017). Concern noted: A descriptive study of editorial expressions of concern in PubMed and PubMed Central. *Research Integrity and Peer Review, 2*(1), 1–11. https://doi.org/10.1186/s41073-017-0030-2

Wilson Sayres, M. A., Hauser, C., Sierk, M., Robic, S., Rosenwald, A. G., Smith, T. M., Triplett, E. W., Williams, J. J., Dinsdale, E., Morgan, W. R., Burnette, J. M., Donovan, S. S., Drew, J. C., Elgin, S. C. R., Fowlks, E. R., Galindo-Gonzalez, S., Goodman, A. L., Grandgenett, N. F., Goller, C. C., & Pauley, M. A. (2018). Bioinformatics core competencies for undergraduate life sciences education. *PLoS One, 13*(6), 1–20. https://doi.org/10.1371/journal.pone.0196878

Yucel, R. (2018). Scientists' ontological and epistemological views about science from the perspective of critical realism. *Science and Education, 27*, 407–433. https://doi.org/10.1007/s11191-018-9983-x