Philosophy of Medicine
Analysis

Diagnostic Parsimony: Ockham Meets Bayes

Bengt Autzen[1]

1 Department of Philosophy, University College Cork, Cork, Munster, Ireland. Email: bengt.autzen@ucc.ie

Abstract

Ockham's razor is the idea that simpler hypotheses are to be preferred over more complex ones. In the context of medical diagnosis, this is taken to mean that when a patient has multiple symptoms, a single diagnosis should be sought that accounts for all the clinical features, rather than attributing a different diagnosis to each. This paper examines whether diagnostic parsimony can be justified by reference to probability theory. I argue that while attempts to offer universal justifications of diagnostic parsimony fail, a more constrained use of this diagnostic principle can be supported.

1. Introduction

Ockham's razor, sometimes referred to as the principle of parsimony, is an inferential principle attributed to the medieval scholar William of Ockham. In its original dictum the principle asserts that entities should not be multiplied unnecessarily in an explanation. This is typically understood as saying that in science simpler hypotheses are to be preferred over more complex ones. Medical diagnosis refers to the identification of a disease, condition, or injury from its signs and symptoms. In the context of medical diagnosis, the principle of parsimony is taken to mean that when a patient has multiple symptoms, a single diagnosis should be sought that accounts for all the clinical features, rather than attributing a different diagnosis to each (Schattner 2015). Diagnostic parsimony recommends choosing the simpler diagnostic hypothesis postulating fewer diseases in the case that two competing diagnostic hypotheses can both account for the clinical data.

While Ockham's razor is widely applied in medical diagnosis, it conflicts with other diagnostic principles invoked by physicians. Hickam's dictum states that multiple symptoms may be due to more than one disease (Hilliard et al. 2004). An example, sometimes referred to as Saint's triad, is the combination of hiatal hernia, gallbladder disease, and diverticulosis. Since there is no pathophysiological basis for the coexistence of these three diseases, Saint's triad seems to demonstrate that more than one disease may be responsible for a patient's symptoms.

To complicate things further, Bayesianism is used as a framework for medical decision making (Sox, Higgins, and Owens 2013). Bayesians assess diagnostic hypotheses based on

ISSN 2692-3963 | Date Received 14-06-2022 | Date Accepted 15-11-2022 | Date First Published 01-12-2022
Philosophy of Medicine | DOI 10.5195/pom.2022.123 | Volume 3 | Issue 1 | pp.1–12

their probabilities. More specifically, given a diagnostic hypothesis H and some set of symptoms S, Bayesians calculate the posterior (or post-test) probability of a diagnostic hypothesis $P(H|S)$ by means of Bayes's theorem, which states:

$$P(H|S) = \frac{P(S|H)P(H)}{P(S)}$$

where $P(H)$ refers to the prior (or pre-test) probability of the diagnostic hypothesis. As the posterior probability approaches 1, the diagnosis becomes nearly certain. Physicians would typically consider the diagnosis confirmed once its posterior probability exceeds a threshold of sufficient certainty (Richardson 2007). Bayesianism also offers an answer to the question of when to act (for example, order a diagnostic test, or prescribe treatment) in the face of diagnostic uncertainty. Bayesian decision theory says that physicians should make treatment decisions based on whether the posterior probability of a diagnosis exceeds a treatment threshold. According to this rule, physicians administer treatment if the probability of disease is above a specified threshold and withhold treatment otherwise. To calculate the thresholds, physicians have to take into account data on the effects of treatments, the diagnostic test's sensitivity and specificity, and the harms of the test (Pauker and Kassirer 1980; Djulbegovic et al. 2015).

This situation naturally leads to the question of how diagnostic parsimony relates to probability theory. One possible answer is that there is no formal relation to be found between different diagnostic principles, such as diagnostic parsimony and Bayesian reasoning. Based on this view, diagnostic principles form a rather loose collective. Ami Schattner (2015), for instance, lists a number of key diagnostic principles, such as Ockham's razor, the law of imperfection, and the rule-out worst-case scenario principle, without explaining how potential trade-offs between different principles can be resolved. While such a view might be the right position to adopt, developments such as the use of artificial intelligence (AI) in medical diagnosis (Fraser, Coiera, and Wong 2018) force us to think more deeply about how to formalize and relate different diagnostic principles. This paper is a step in that direction.

In this paper I examine the relationship between diagnostic parsimony and probabilistic reasoning. A first proposal, here referred to as the "universal approach," suggests that diagnostic parsimony is directly implied by the probability calculus. Based on this view, a single diagnosis that accounts for all symptoms is to be preferred over attributing a different diagnosis to each symptom *independently* of the precise content of the simple and complex diagnostic hypotheses under consideration. A second proposal, defended here and referred to as the "domain-specific approach," only offers a much weaker form of support for diagnostic parsimony. The domain-specific approach asserts that in some clinical applications diagnostic parsimony offers a valid principle for differential diagnosis in the sense that it agrees with Bayesian principles but fails to identify the diagnostic hypothesis with the highest posterior probability in other contexts. According to the domain-specific reading, the precise content of the competing simple and complex diagnostic hypotheses matters for assessing the validity of diagnostic parsimony. The agreement between diagnostic parsimony and Bayesian reasoning is contingent on empirical facts about the world and not a mathematical certainty. Or, to put it more bluntly, sometimes the simpler hypothesis does not adequately account for the complexity of the world. Importantly, both the universal and the domain-specific approaches aim to derive diagnostic parsimony from Bayesian principles. Therefore, Bayesian epistemology is considered more fundamental than diagnostic parsimony.

At this stage a comment on methodology is in order. This paper is concerned with normative aspects of medical decision making. It is widely held that Bayesianism (or expected utility theory) offers the correct normative standard for decision making under uncertainty (Savage 1954). This does not imply, however, that Bayesianism offers a correct descriptive theory of human decision making under uncertainty. Put differently, for the purpose of this paper, it is inconsequential that the kind of numerical calculations required by Bayesian decision theory are not regularly performed by physicians in medical practice. This paper engages in an ideal theory of medical decision making; it examines the epistemic status of diagnostic parsimony for an ideal Bayesian reasoner.[1]

The structure of this paper is as follows: In section 2 I argue that the universal approach to justifying diagnostic parsimony is flawed. In particular, I point out that the reasoning underlying this justification leads to some unacceptable diagnostic conclusions. In contrast, I argue for the weaker domain-specific approach in section 3. While the latter view offers only limited support for diagnostic parsimony as a valid diagnostic principle, it resolves the need to justify apparent violations of the principle of diagnostic parsimony. The view defended here assigns diagnostic parsimony the status of a heuristic decision tool that sometimes agrees with more fundamental probabilistic considerations but should not by itself override probabilistic reasoning. More specifically, I suggest that diagnostic parsimony can serve as a heuristic when encountering young and otherwise healthy patients in general practice. However, this diagnostic principle is problematic in fields such as geriatrics when dealing with chronically ill patients with multiple morbidities. In section 4 I discuss (and resolve) some further counterexamples to diagnostic parsimony suggested in the medical literature. In section 5 I relate my analysis of diagnostic parsimony to recent discussions of Ockham's razor in the philosophical literature. Section 5 is primarily directed at a philosophical audience and can be skipped by the medical reader. In section 6 I conclude with some final comments.

2. The Universal Approach

In the statistical literature it has been argued that Bayesian inference will automatically assign greater likelihood to a simpler hypothesis if the data are compatible with both a simpler and a more complex hypothesis (Jefferys and Berger 1992). David MacKay, for instance, asserts that "coherent inference (as embodied by Bayesian probability) automatically embodies Occam's razor, quantitatively" (2003, 344).[2] Turning to medical diagnosis, an example of the idea that Ockham's razor falls directly out of the probability calculus is provided by Harold C. Sox, Michael C. Higgins, and Douglas K. Owens, who claim that diagnostic parsimony "is based on a basic theorem of probability theory" (2013, 19). More specifically, they write:

---

[1] For a further discussion of the relationship between normative and descriptive decision theory in medical decision making, see Djulbegovic et al. (2015).

[2] For a philosophical discussion of what is referred to as "Bayesian Ockham's Razor" in the statistical literature, see Sober (2015).

> The probability that two unrelated events occur simultaneously is the probability of one event multiplied by the probability of the other event. The product of the two probabilities is a much lower number than the probability of either event occurring by itself. The rule of parsimony is probably reliable in previously healthy people but may be less reliable in persons with several chronic diseases. In this case, two diagnostic hypotheses are less likely to be independent of each other, which increases the probability that both are present. (Sox, Higgins, and Owens 2013, 19)

To illustrate, consider a person that presents with symptoms consistent with bacterial pneumonia but has also started coughing up small amounts of blood.[3] Coughing up blood (hemoptysis) is an unusual, though well-known, complication of bacterial pneumonia. However, coughing up blood is also associated with having an airway tumor. In order to account for these observations, two possible explanations are considered: the simple (or unifying) diagnosis asserts that the patient has bacterial pneumonia; the complex (or disunifying) diagnosis asserts that the patient has bacterial pneumonia *and* an airway tumor.

One might wonder why more possible explanations of the patient's symptoms are not being considered in the example. For instance, one might also suspect that the patient suffers from tuberculosis infection of the lung. The reason is that an example involving only two distinct diseases seems to offer the best prospect for providing a Bayesian justification of Ockham's razor in a diagnostic context. I return to the issue of using Ockham's razor for diagnosing more possible diseases in section 4. A clinician might also ask for more details on the patient's symptoms than provided in the example. For instance, how many times has the patient been coughing during the day? Is the cough worse at night? Has the cough worsened over the past two days? While further details of the patient's symptoms would improve the realism of the example, the litmus test for this idealized ("toy") example is to probe Sox, Higgins, and Owens's proposed relationship between probability theory and diagnostic parsimony. It is not implied that a medical doctor will reach a final diagnosis based on the limited information provided in the example. Indeed, a doctor will, in all likelihood, contemplate further testing, such as ordering a chest X-ray, a chest CT scan, a sputum culture, a pulmonary arteriography, or a lung biopsy. I explore the relationship between probability theory, Ockham's razor, and further testing in section 3. The question I pursue here is which of the two hypothesized diagnoses is more probable in light of the available evidence.

Returning to the example, let us refer to a patient having bacterial pneumonia as event D_1 and a patient having an airway tumor as event D_2. Now, Sox, Higgins, and Owens seem to suggest that there is a probabilistic argument to the effect that the simple diagnostic hypothesis that the patient has bacterial pneumonia has higher probability than the complex hypothesis that the patient has both bacterial pneumonia and an airway tumor. If the two diagnostic hypotheses D_1 and D_2 are probabilistically independent, the probability of the joint event $D_1 \cap D_2$ equals the product of the probabilities $P(D_1)$ and $P(D_2)$. Assuming that both probabilities—$P(D_1)$ and $P(D_2)$—are strictly larger than zero and strictly smaller than 1, then the product of the two probabilities is a much lower number than the probability of event D_1 occurring by itself. So, assuming that bacterial pneumonia and an airway tumor

[3] I would like to thank Mark Tonelli for suggesting the example.

are probabilistically independent for this patient, one has established a formal argument to the effect that the simple diagnosis has higher probability than the more complex one.

Probabilistic independence, however, is a red herring in this context. It follows directly from the axioms of probability theory that the probability of the joint event $D_1 \cap D_2$ can never be strictly larger than the probability of the single event D_1. This result holds whether D_1 and D_2 are probabilistically independent or not. As a consequence, the simple diagnostic hypothesis is always at least as probable as the more complex diagnostic hypothesis. This observation seems to offer a strong—indeed, I will argue "too strong"—justification for diagnostic parsimony. Independent of whether the patient is an otherwise healthy teenage girl, or a 70-year-old man with a long history of smoking, the diagnosis of bacterial pneumonia is always at least as probable as the hypothesis that the patient has both bacterial pneumonia and an airway tumor. If you only care about the most probable diagnostic hypothesis, there is never a principled reason to opt for the more complex diagnosis. Parsimony rules.

Something has gone wrong. I will argue that the problem is to be found in the way we represent the simple diagnostic hypothesis under consideration. So far, we represented the simple diagnosis that the patient has bacterial pneumonia as diagnostic hypothesis $H_1 = D_1$. Indeed, this representation naturally lends itself to Sox, Higgins, and Owens's argument comparing the probability of a single event (or hypothesis) to the product of the probabilities of two events (or hypotheses). However, the diagnostic hypothesis H_1 is indifferent as to whether the patient also has an airway tumor. More formally, the hypothesis $H_1 = D_1$ is logically equivalent to the hypothesis $(D_1 \cap D_2) \cup (D_1 \cap \neg D_2)$; that is, the hypothesis H_1 refers to the situation that the patient either has bacterial pneumonia and an airway tumor $D_1 \cap D_2$ or bacterial pneumonia but no airway tumor $D_1 \cap \neg D_2$.

What clinicians seem to have in mind when diagnosing bacterial pneumonia is that the patient has only bacterial pneumonia but not also an airway tumor. The simpler hypothesis of bacterial pneumonia is therefore adequately represented by means of the diagnostic hypothesis $D_1 \cap \neg D_2$. This different representation of the hypothesis in question has important consequences for a formal justification of diagnostic parsimony. Both Sox, Higgins, and Owens's argument invoking probabilistic independence and the argument invoking the monotonicity of the probability measure rely on representing the simple diagnostic hypothesis by means of a single hypothesis while the more complex hypothesis is represented by means of a conjunction that involves the simple hypothesis and another one. This line of reasoning and, hence, this justification of diagnostic parsimony is blocked when changing the representation of the simple hypothesis from D_1 to $D_1 \cap \neg D_2$.

3. The Domain-Specific Approach

Where does this leave the epistemic status of diagnostic parsimony? In the previous section I demonstrated that a universal application of this principle cannot be recommended. In this section I suggest that only a more limited use of diagnostic parsimony can be supported. A simple diagnostic hypothesis can have higher probability than a more complex hypothesis in one context but not in others; the assessment depends on what the simpler diagnostic hypothesis H_1 and the more complex diagnosis H_2 stand for. As a consequence, there is no truly general justification of diagnostic parsimony that is independent of the precise empirical content of the particular diagnostic hypotheses under consideration. Diagnostic parsimony is only warranted in a domain-specific way—that is, depending on the diagnostic context at hand.

Let us return to the patient presenting with symptoms consistent with bacterial pneumonia and who also coughs up blood. Again, consider the two possible diagnostic hypotheses $D_1 \cap D_2$ and $D_1 \cap \neg D_2$. In order to assess the relative merits of these two diagnoses from a Bayesian perspective, one has to compare the posterior probabilities $P(D_1 \cap D_2|S)$ and $P(D_1 \cap \neg D_2|S)$, with S denoting the whole set of symptoms shown by the patient. Comparing the numerical values of the probabilities $P(D_1 \cap D_2|S)$ and $P(D_1 \cap \neg D_2|S)$ amounts to comparing the two products $P(S|D_1 \cap D_2) * P(D_1 \cap D_2)$ and $P(S|D_1 \cap \neg D_2) * P(D_1 \cap \neg D_2)$. Since coughing up blood is seen as a less common symptom of bacterial pneumonia, it is assumed here that the likelihood $P(S|D_1 \cap D_2)$ is, at least somewhat, larger than the likelihood $P(S|D_1 \cap \neg D_2)$; that is, it is more probable to observe the whole set of symptoms if the patient has both bacterial pneumonia and an airway tumor than if the patient has only bacterial pneumonia.

Having compared the likelihoods of the two competing diagnostic hypotheses, one has to assess their prior probabilities in a next step. What is the prior probability of having only bacterial pneumonia as opposed to having bacterial pneumonia and an airway tumor? Here, it seems crucial to learn more about the medical background of the patient under consideration in order to address this question. It is plausible to assume that the prior probability of a young and otherwise healthy patient having both bacterial pneumonia and an airway tumor is extremely low, while this probability significantly increases in an elderly patient with a history of smoking.

Returning to our quantitative assessment of the two diagnostic hypotheses, consider first a young (and otherwise healthy) patient. While the likelihoods slightly favor the more complex hypothesis that the patient has both bacterial pneumonia and an airway tumor, the prior probabilities speak clearly in favor of the simpler hypothesis that the patient only has bacterial pneumonia. Assuming that the effect of the lower prior probability of the complex hypothesis counteracts its higher likelihood, the comparison of the posterior probabilities of the two hypotheses will favor the simpler hypothesis. This line of reasoning provides some support for using diagnostic parsimony as a heuristic tool in general practice when encountering young and otherwise healthy patients. Heuristics can assist our reasoning in complex scenarios but ultimately sound probabilistic reasoning should support our diagnoses.

Turning to an elderly patient with a history of smoking, the likelihood assessment remains unchanged. Again, the more complex hypothesis postulating both the presence of bacterial pneumonia and an airway tumor scores somewhat better when only the likelihoods are considered. With regard to the prior probabilities of the two diagnostic hypotheses, it is plausible to assume that there is not such a dramatic numerical difference between $P(D_1 \cap D_2)$ and $P(D_1 \cap \neg D_2)$ as it exists in the young patient. As a consequence, it becomes an open question whether the simpler or the more complex diagnostic hypothesis will be favored when comparing the products of the likelihood and prior probability of the two hypotheses. In any case, it is much less clear that the simpler diagnosis of bacterial pneumonia will come out on top for this patient in the Bayesian assessment. This analysis suggests that the use of diagnostic parsimony as a heuristic is problematic in fields such as geriatrics, where the clinician frequently encounters chronically ill patients with multiple morbidities.

So far, the discussion has focused on the question of which of two possible diagnoses has higher probability in the light of some observed symptoms. It is important to note that establishing that a diagnosis has a higher probability than a competitor does not imply that further diagnostic testing is to be ordered, or treatment for the condition to be initiated. Stephen G. Pauker and Jerome P. Kassirer provide criteria for making theses medical decisions within a Bayesian (or expected utility) framework (1975, 1980). They examine when a physician should prescribe treatment under the assumption that no further diagnostic information is available (1975). They establish a treatment threshold; that is, the value of the probability of a patient having disease where one is indifferent between treating and not treating. The threshold T is given by:

$$T = \frac{1}{1+\frac{\Delta B}{\Delta H}},$$

where ΔB denotes the net benefit—the difference in the expected utility of the outcomes if the diseased patient were treated or not treated—and ΔH is the net harm—the expected difference in the utility of the outcome if not treated versus if treated for a patient without the disease. The rational course of action for a clinician is to apply treatment if the probability of the disease exceeds threshold T and to withhold treatment if the probability of disease is smaller than this threshold. Having established that a simple diagnosis has a higher probability than its more complex competitor does not imply that the probability of the simple diagnosis exceeds threshold T. Doing so will depend on the numerical values of the parameters entering the calculation of threshold T. So, while there is a meaningful way of talking about diagnostic parsimony in terms of diagnostic hypotheses having higher probability, there is a parting of ways between diagnostic parsimony and medical decision making.

The same applies when more complex Bayesian models of medical decision making are considered. Pauker and Kassirer (1980) discuss the question of whether a clinician should: (a) continue observing the patient without treatment or testing; (b) perform a diagnostic test and subsequently base the treatment on the test result; or (c) administer treatment without requesting a test. They identify two thresholds for medical decision making: a testing threshold and test-treatment threshold. Based on the model, the clinician should treat the patient without further testing if the posterior probability of the disease is larger than the testing threshold but should continue to observe the patient without ordering a test or treatment if the probability is smaller than the test-treatment threshold. The clinician should request a test if the probability of the disease is between the testing and the test-treatment thresholds. Again, having established that a simpler diagnosis has higher probability than a more complex diagnosis, based on a set of observed symptoms, does not imply whether a threshold has been reached and, if so, which one. Diagnostic parsimony is an evidential principle not directly linked to medical decision making, as the latter requires more than information about the posterior probabilities of diagnostic hypotheses.

The domain-specific approach outlined above sits fairly well with views expressed in some recent exchanges on diagnostic parsimony in the medical literature. Commenting on a survey by Schattner (2015) that lists key principles for teaching clinical medicine, including the law of parsimony, Oscar Jolobe (2016) argues that in some medical contexts diagnostic parsimony should be rejected. He refers to the example of immunocompromised patients, such as those with HIV and AIDS, in which bacterial pneumonia can coexist with active pulmonary tuberculosis. In response, Schattner (2016) acknowledges that active tuberculosis can coexist with other infections and, as such, constitutes a counterexample to diagnostic parsimony. More generally, while Schattner considers diagnostic parsimony an important concept in medicine, he accepts that exceptions to the principle are rather abundant.

If diagnostic parsimony merely constitutes a heuristic for medical decision making that is based on empirical regularities about relative frequencies of diseases and their symptoms observed in medical practice, it should not come as a surprise that there can be exceptions to the "law" of parsimony. Indeed, the view outlined in this paper resolves the need to justify apparent violations of diagnostic parsimony. Matters would be different if there existed a universal justification for diagnostic parsimony based on the probability calculus. In that case, any exception to the principle of parsimony would be on pain of violating the axioms of probability that are widely held to embody the principles of rational reasoning under uncertainty.

4. Further Challenges to Diagnostic Parsimony

So far, the discussion has focused on diagnostic hypotheses involving two distinct diseases, bacterial pneumonia and an airway tumor. This can be considered as the most promising scenario for providing a universal justification of diagnostic parsimony. Once three different diseases are considered, the prospect of a universal justification looks more bleak. Indeed, Sox, Higgins, and Owens (2013) point out that having two common diseases can be more probable than having a single but rare disease. Marta Freixa et al. (2019) go one step further by describing a case of two rare diseases occurring in a patient, which they consider to be in violation of diagnostic parsimony.

In this section I contrast these two scenarios and argue that while the case of comparing two common diseases with one single rare disease poses a challenge for diagnostic parsimony, the second case of comparing two rare diseases with a single rare disease does not. Let us consider three diseases: D_1, D_2, and D_3. To begin with, assume that D_1 is a rare disease whereas D_2 and D_3 are both common diseases. Now, consider the two diagnostic hypotheses $H_1 = D_1 \cap \neg D_2 \cap \neg D_3$ and $H_2 = \neg D_1 \cap D_2 \cap D_3$ and suppose that the likelihoods $P(S|H_1)$ and $P(S|H_2)$ are approximately equal, where S denotes the set of symptoms of the patient. Both hypotheses H_1 and H_2 account for the observed symptoms reasonably well. Under these assumptions, it is plausible to assume that the more complex hypothesis H_2 has a higher posterior probability than the simple hypothesis H_1. If diseases D_1, D_2, and D_3 are probabilistically independent, this would require that the product $P(\neg D_1) * P(D_2) * P(D_3)$ exceeds the product $P(D_1) * P(\neg D_2) * P(\neg D_3)$. In that case, a physician opting for the most probable diagnosis should choose the more complex diagnosis H_2 instead of the simpler hypothesis H_1. Such a diagnostic decision is in conflict with diagnostic parsimony; it constitutes a systematic violation of this principle.

In the second example, we assume that the diseases D_2 and D_3 are rare, rather than common. Again, we presume that the likelihoods of diagnoses $H_1 = D_1 \cap \neg D_2 \cap \neg D_3$ and $H_2 = \neg D_1 \cap D_2 \cap D_3$ are approximately equal, and that the diseases D_1, D_2, and D_3 are probabilistically independent. In that case, it is plausible to assume that the simpler hypothesis H_1 has the larger posterior probability than the more complex hypothesis H_2 due to the product of prior probabilities $P(\neg D_1) * P(D_2) * P(D_3)$ being smaller than the product $P(D_1) * P(\neg D_2) * P(\neg D_3)$. As a consequence, a Bayesian clinician opting for the most probable diagnosis would prefer the simpler diagnosis H_1 over the more complex hypothesis H_2 in line with diagnostic parsimony.

Now, Freixa et al. seem to suggest that the case of a patient with two rare tumors challenges diagnostic parsimony and speaks in favor of Hickam's dictum. However, they do not discuss an alternative diagnostic hypothesis invoking only one disease in competition with the diagnosis that the patient has two rare tumors simultaneously. This is problematic if their example is considered a case in favor of Hickam's dictum and against Ockham's razor. Diagnostic parsimony presumes that there are two diagnoses, a simple one and a complex one, which are both consistent with the symptoms. I therefore resist the conclusion that the case report speaks against diagnostic parsimony. Rather, the example demonstrates that improbable events happen. This in itself, however, does not invalidate an inference principle such as diagnostic parsimony. Inductive inference is not foolproof.

Consider an analogy. It is very improbable that a person will win the lottery twice but it can (and does) happen. However, this does not invalidate the view that when forming beliefs about the world in the light of our evidence, one should opt for a more probable alternative, such as the hypothesis that one does not win the lottery twice. Similarly, the patient with two rare and probabilistically independent tumors has "won" twice in a lottery with very unfortunate outcomes. When deciding on the most plausible diagnosis, opting for the most probable diagnostic hypothesis looks like a rational choice. This view is not undermined by the fact that sometimes improbable things happen.

Matters could, of course, be different. Again, assume that there is a single disease D_1 that is consistent with the symptoms in Freixa et al.'s case report. Now, suppose further that the likelihood of this disease given the symptoms S is very small. Even though disease D_1 can account for all the observed symptoms observed in the patient, these symptoms are very uncommon given disease D_1. In that case, it could well be that the posterior probability of diagnosis H_2 invoking the two diseases D_2 and D_3 exceeds the posterior probability of the simple diagnosis H_1. The most probable diagnosis then postulates that the patient has two rare tumors. Hickam's dictum would be supported by Bayesian reasoning.

5. Philosophical Perspectives on Parsimony

It is instructive to relate the present discussion to some prominent views on parsimony found in the philosophical literature. Elliott Sober (2015) makes the case for a reductionist view on parsimony, according to which Ockham's razor is epistemically relevant only insofar as it contributes to the achievement of some more fundamental epistemic goal.[4] For instance, Bayesians primarily care about the posterior probabilities of hypotheses; considerations of parsimony are only derived from these primary considerations. Both the universal and the domain-specific approach of justifying diagnostic parsimony are examples of a reductionist view on parsimony. Take the domain-specific approach: by saying that, in certain contexts, diagnostic parsimony can reliably assist in identifying the diagnostic hypothesis with the higher posterior probability, parsimony is epistemically

[4] For an early statement of Sober's reductionist view on parsimony, see Sober (1990).

relevant insofar as it contributes to the epistemic goals of a Bayesian clinician. The same applies to the universal approach.

The previous discussion suggests a distinction with regard to the kind of reductive relationship between parsimony and Bayesianism. Reductionism about parsimony requires that the epistemic value of Ockham's razor is derived from some more fundamental epistemic principles, such as identifying hypotheses with high posterior probability. Now, it could be that Ockham's razor *always* serves this more fundamental epistemic goal, or that it does so only in *certain* situations. The former scenario might be referred to as universal reductionism; the latter might be called domain-specific reductionism. While in the former case, the recommendations of Ockham's razor are always in line with the Bayesian analysis, in the latter case the two can part ways. If the principle of parsimony was implied by a theorem of probability theory, as suggested by the universal approach, a Bayesian clinician could not avoid applying Ockham's razor as long as she correctly computes the probabilities figuring in Bayes's theorem. The picture that emerges from the present discussion of diagnostic parsimony, however, suggests otherwise. Diagnostic parsimony and a Bayesian analysis of diagnostic hypotheses can part ways. Clinicians applying diagnostic parsimony and clinicians adopting a Bayesian framework can come to different diagnoses.

In his earlier writings, Sober distinguishes between parsimony being a "local" and a "global" virtue in scientific reasoning (1988). Parsimony, understood as a global epistemic virtue, applies to all fields of human inquiry and means the same thing across these domains. Put simply, the principle of parsimony is part and parcel of good scientific inference. Traditionally, this view on parsimony has figured prominently in the philosophical discussion—for example, Popper (1959); Kemeny (1953); and Goodman (1958). In contrast, parsimony seen as a local epistemic virtue means one thing in one domain while it may mean something different in another domain. Additionally, whether a more parsimonious hypothesis is to be preferred is also a local matter. Based on the latter reading, Ockham's razor is not a universally valid rule of scientific inference.

The contrast between global and local parsimony clearly sits well with the distinction between universal and domain-specific reductionism. In particular, the domain-specific approach to diagnostic parsimony instantiates Sober's reductive, local understanding of parsimony in the context of medical diagnosis. It is worth noting, however, that parsimony could be a global epistemic virtue but the universal approach to diagnostic parsimony, as understood here, could still not hold. For instance, the principle of parsimony could be a central principle in scientific reasoning without being closely linked to Bayesianism for the simple reason that Bayesianism does not provide the right framework for scientific inference—an idea that might seem attractive to non-Bayesian philosophers of science such as Deborah G. Mayo (1996). Sober's distinction between local and global parsimony is applied by Anya Plutynski (2005) to shed light on the Fisher-Wright debate in evolutionary biology. Plutynski defends the view that parsimony is a local epistemic virtue and argues that some participants in the Fisher-Wright debate in effect appeal to facts about the prior probability or likelihood of scientific hypotheses when they claim that Fisher's theory is more parsimonious than its competitor. The present paper can be seen as a further example illustrating the local epistemic virtue of parsimony in a particular scientific application—here, medical diagnosis.

6. Conclusion

This paper has examined diagnostic parsimony; that is, a medical application of Ockham's razor. I have argued two points, one negative and one positive. First, I argued that diagnostic parsimony is not implied by the probability calculus. When thought through, this approach will lead to some unacceptable diagnostic conclusions by always prioritizing simpler diagnostic hypotheses. Second, I offered some support for what I called a domain-specific approach to justifying diagnostic parsimony. According to this view, diagnostic parsimony can serve as a heuristic for medical diagnosis to the Bayesian clinician in some but not all applications. This view coincides with the local, reductionist reading of parsimony initially defended by Sober in the philosophical literature. The domain-specific approach to diagnostic parsimony instantiates the view that parsimony is a local epistemic virtue in scientific reasoning.

Acknowledgments

Funding from the Austrian Science Fund (Lise Meitner Grant M2366-G24) is gratefully acknowledged.

Disclosure Statement

No competing interest was reported by the author.

References

Djulbegovic, Benjamin, Jef van den Ende, Robert M. Hamm, Thomas Mayrhofer, Iztok Hozo, and Stephen G. Pauker. 2015. "When Is Rational to Order a Diagnostic Test, or Prescribe Treatment: The Threshold Model as an Explanation of Practice Variation." *European Journal of Clinical Investigation* 45, no. 5: 485–493. https://doi.org/10.1111/eci.12421.

Fraser, Hamish, Enrico Coiera, and David Wong. 2018. "Safety of Patient-Facing Digital Symptom Checkers." *The Lancet* 392, no. 10161: 2263–2264. https://doi.org/10.1016/s0140-6736(18)32819-8.

Freixa, Marta, André F. Simões, Joana Barbosa Rodrigues, Sara Úria, and Glória Nunes da Silva. 2019. "Occam's Razor versus Hickam's Dictum: Two Very Rare Tumours in One Single Patient." *Oxford Medical Case Reports* 2019, no. 5: 195–198. https://doi.org/10.1093/omcr/omz029.

Goodman, Nelson. 1958. "The Test of Simplicity." *Science* 128, no. 3331: 1064–1069. https://doi.org/10.1126/science.128.3331.1064.

Hilliard, Anthony A., Steven E. Weinberger, Lawrence M. Tierney, David E. Midthun, and Sanjay Saint. 2004. "Occam's Razor versus Saint's Triad." *New England Journal of Medicine* 350, no. 6: 599–603. https://doi.org/10.1056/nejmcps031794.

Jefferys, William H. and James O. Berger. 1992. "Ockham's Razor and Bayesian Analysis." *American Scientist* 80, no. 1: 64–72. https://www.jstor.org/stable/29774559.

Jolobe, Oscar. 2016. "Caveats to the Application of Diagnostic Parsimony." *QJM* 109, no. 2: 143. https://doi.org/10.1093/qjmed/hcv157.

Kemeny, John G. 1953. "The Use of Simplicity in Induction." *Philosophical Review* 62, no. 3: 391–408. https://doi.org/10.2307/2182878.

MacKay, David. 2003. *Information Theory, Inference and Learning Algorithms*. Cambridge: Cambridge University Press.

Mayo, Deborah G. 1996. *Error and the Growth of Experimental Knowledge*. Chicago: University of Chicago Press.

Pauker, Stephen G. and Jerome P. Kassirer. 1975. "Therapeutic Decision Making: A Cost Benefit Analysis." *New England Journal of Medicine* 293, no. 5: 229–234. https://doi.org/10.1056/nejm197507312930505.

———. 1980. "The Threshold Approach to Clinical Decision Making." *New England Journal of Medicine* 302, no. 20: 1109–1117. https://doi.org/10.1056/NEJM198005153022003.

Plutynski, Anya. 2005. "Parsimony and the Fisher-Wright Debate." *Biology and Philosophy* 20, no. 4: 697–713. http://dx.doi.org/10.1007/s10539-004-7044-0.

Popper, Karl R. 1959. *The Logic of Scientific Discovery*. London: Hutchinson.

Richardson, W. Scott. 2007. "We Should Overcome the Barriers to Evidence-Based Clinical Diagnosis." *Journal of Clinical Epidemiology* 60, no. 3: 217–227. https://doi.org/10.1016/j.jclinepi.2006.06.002.

Savage, Leonard J. 1954. *The Foundations of Statistics*. New York: John Wiley & Sons.

Schattner, Ami. 2015. "Teaching Clinical Medicine: The Key Principals." *QJM* 108, no. 6: 435–442. http://dx.doi.org/10.1093/qjmed/hcv022.

———. 2016. "Caveats to the Application of Diagnostic Parsimony—Reply." *QJM* 109, no. 2: 144. http://dx.doi.org/10.1093/qjmed/hcv161.

Sober, Elliott. 1988. *Reconstructing the Past: Parsimony, Evolution and Inference*. Cambridge, MA: MIT Press.

———. 1990. "Explanation in Biology: Let's Razor Ockham's Razor." *Royal Institute of Philosophy Supplement* 27: 73–93. http://dx.doi.org/10.1017/S1358246100005051.

———. 2015. *Ockham's Razors: A User's Manual*. Cambridge: Cambridge University Press.

Sox, Harold C., Michael C. Higgins, and Douglas K. Owens. 2013. *Medical Decision Making*. Oxford: Wiley-Blackwell.