Word Count (including references and notes): 10,412 [without = 6628]

Probability and Informed Consent

Nir Ben-Moshe
University of Illinois at Urbana-Champaign

Ben Levinstein
University of Illinois at Urbana-Champaign

Jonathan Livengood
University of Illinois at Urbana-Champaign

Abstract. We argue that securing informed consent requires not only that patients understand the probabilities of various risks and benefits of proposed treatments or procedures they face but also that securing informed consent requires communicating how probability expressions (or other tools for representing uncertainty) are to be interpreted, the quality and quantity of the evidence for the probabilities reported, and how these probability claims might or might not be relevant to a patient's decisions. We conclude by considering two possibilities. Either patients cannot understand the difficult concepts and issues we discuss in this paper and so cannot be genuinely informed of their risks, or patients can come to understand the relevant issues when properly advised, such that the issues we discuss are not a principled barrier to obtaining informed consent. If patients cannot understand the relevant issues, then the informed consent requirement must be relaxed so as not to include a requirement of reporting probability claims to patients. If patients can come to understand the relevant issues when properly advised, then the medical community should train physicians to provide that advice and help patients understand the advice (potentially through third-party patient activist decision theorists who are members of broader patient activist groups).

Keywords: Informed Consent, Probability, Bayesianism, Frequentism, Decisions

Acknowledgements: [[Kindnesses go here.]]

Probability and Informed Consent

Medical interventions do not always succeed. Sometimes they cause serious health problems or even death. Since the outcome of any given medical intervention is uncertain, decisions about which, if any, to choose for oneself or to recommend to a patient should be sensitive to that uncertainty. Hence, understanding and communicating uncertainty is crucially important for medical professionals. Probability theory and statistics provide one natural, widely-used tool for representing uncertainty and for drawing uncertain inferences.[1]

We think it is advisable for physicians to use probability and statistics to communicate to their patients what is known about the risks and benefits associated with available procedures, though we do not take any position with respect to *best practices* for successfully communicating first-order uncertainty in its own right.[2] In this paper, we argue that securing informed consent

[1] Probability theory and statistics are not the *only* tools one might use to represent and communicate uncertainty. However, the issues we want to raise also matter for alternatives, so we will focus on probability theory and statistics in this paper.

[2] We are neutral with respect to whether one should use verbal probability expressions (such as 'likely' or 'probable') or numerical probability expressions (such as '20-percent chance'). Early studies showed that there is some degree of consistency in verbal expressions of probability (Kong et al. 1986) but that this consistency is rather limited (Mapes 1979). Some empirical research suggests that physicians assign a wide range of numerical values to subjective verbal expressions of probability and that the choice of verbal probability expressions is influenced by the severity of the consequences associated with the risks (Merz et al. 1991). A meta-analysis on the role of translations of verbal into numerical probability expressions in risk management found that ratings of verbal probability statements varied considerably: at the extreme, differences as large as 28 percentage points were observed for the very commonly used term 'possible' (Theil 2002). A more recent study showed that while psychiatrists communicate probabilities "verbally," there is considerable variation in their associated numerical estimates (describing something as happening "often" can mean an absolute risk of 10% or 90%, depending on the individual physician), and there may be a comparable amount of heterogeneity in the patients' perceptions too, making a congruence between the doctor's and the patient's perceptions unlikely (Hamann et al. 2011). For recent literature on the use of verbal and numerical probability expressions, see Christopher et al. (2010), Collins and Hahn (2018), Hanauer et al. (2012), Jenkins et al. (2018a), Karelitz and Budescu (2004), McDowell and Jacobs (2017), Spiegelhalter (2017), Tait et al. (2010), and Vahabi (2010).

We are also neutral with respect to a number of related issues having to do with the presentation of probabilistic information. We are neutral with respect to whether one should use a point estimate or a probability range, for which see Dieckmann et al (2010), Jenkins et al. (2018b), Longman et al. (2012), and Sladakovic et al. (2016). We are neutral with respect to whether one should use symbolic-algebraic representations or iconic-geometric representations, for which, see Lipkus (2007), Spiegelhalter et al. (2011), and Zipkin et al. (2014). And we are neutral with respect to what features of patients, such as numeracy (Nelson et al. 2008), affect their

requires more than successfully conveying uncertainty regarding risks and benefits of a proposed treatment or procedure. Securing informed consent also requires communicating how probability expressions (or other tools for representing uncertainty) are to be interpreted, the quality and quantity of the evidence for the probability expressions, and how these probability claims might or might not be relevant to a patient's decisions.

What lesson should we draw? Either patients cannot understand the difficult concepts and issues we discuss in this paper and so cannot be genuinely informed of their risks, or patients can come to understand the relevant issues when properly advised, such that the issues we discuss are not a principled barrier to obtaining informed consent. If patients cannot understand the relevant issues, then the informed consent requirement must be relaxed so as not to include a requirement of reporting probability claims to patients. If patients can come to understand the relevant issues when properly advised, then the medical community should train physicians to provide that advice and help patients understand the advice (potentially through third-party patient activist decision theorists who are members of broader patient activist groups).

We proceed as follows. In Section 1, we briefly present the current state of the literature on informed patient consent with emphasis on disclosing probabilities. In Section 2, we argue that securing informed consent requires significantly more than simply reporting probabilities. We argue in Section 2.1 that when physicians report probabilities, they also need to explain how the concept of probability that they are using works. In other words, they need to specify the way in which they interpret bare probability claims--whether in terms of degrees of belief or relative frequency or something else. We argue in Section 2.2 that patients also must be informed as to

understanding of probability reports. The issues we raise in this paper arise regardless of how these incredibly important and interesting debates are ultimately settled.

the quantity and quality of the evidence for reported probabilities. And we argue in Section 2.3 that patients must be informed regarding how probabilities might or might not be relevant to their decisions. Finally, in Section 3, we step back and reflect on the consequences of our arguments for medical practice.

1. The Nature and Importance of Informed Patient Consent

Many researchers in biomedical ethics recognize the importance of disclosing uncertainty in obtaining informed consent. According to Brock (1987, 121), any attempt to obtain valid informed consent should include the presentation of the following information to the patient: the patient's current medical condition, including a prognosis if no treatment is pursued, treatment alternatives that might improve the patient's prognosis, including explanation of the procedures, the significant risks and benefits of the alternatives with their associated probabilities, and a recommendation of the best alternative.[3] Similarly, Katz (1984, 2) argues that patients need to know "the nature of the ailment, the nature of the proposed treatment, the probability of success or of alternatives, and perhaps the risks of unfortunate results and unforeseen conditions within the body."[4] Young (2009, 536) notes that "patients do need to know about the kind of risks they face, how likely it is that those risks will eventuate, and, if they do, what the effects will be and when they will occur." And Berg et al. (2001, 56-57) argue that there are four elements of the risk that should be disclosed: the nature of the risk, the magnitude of the risk, the probability that the risk might materialize, and the imminence of risk materialization. All of these authors take

[3] Comprehension of the information disclosed is also an important component of informed consent. Medical practitioners should help their patients, as much as they can, comprehend the information they disclose through effective communication (Beauchamp and Childress 2008, ch. 4; Manson and O'Neill 2007).
[4] Katz is quoting from Natanson v. Kline, 350 P. 2d 1093 (Kan. 1960).

the disclosure of the uncertainty regarding prospective procedures to be very important, and most are explicit about including the *probability* of the risks and benefits of various treatment options in order to secure informed consent. Before discussing some problems related to the disclosure of probabilities, we wish to commence by explaining why disclosing probabilities is *prima facie* essential for obtaining informed consent. We do so by appealing to various justifications for requiring that physicians obtain informed consent from their patients.[5]

The dominant justification for informed consent provided in the literature is respect for autonomy (Faden and Beauchamp 1986; Beauchamp and Childress 2008, ch. 4; Beauchamp 2010).[6] Recently, some authors have focused instead on the role informed consent plays in regulating how patients may waive their ethical and legal rights (Manson and O'Neill 2007, 72-77), or more generally, in governing the way normative boundaries between patient and medical practitioner may be drawn (Millum and Bromwich, 2018). Other justifications for the practice of requiring informed patient consent include preventing bodily trespass and promoting self-ownership (Archard, 2008), preventing abusive conduct (Millum and Bromwich, 2013), and building trust in the medical profession (O'Neill 2002, ch. 7).[7] The justification from autonomy requires that medical practitioners provide the patient with information that allows the patient to deliberate and genuinely choose a course of action for herself. Respect for patient autonomy also requires practitioners to do their best to provide information that allows the patient to choose a

[5] In this paper, we are concerned with informed consent in the physician-patient relationship, but what we say also applies in biomedical research settings.

[6] Patient autonomy as a justification for informed consent has been endorsed by the National Commission for the Protection of Human Subjects of Biomedical and Behavioral Research, 1979, B(1) & C(1), and the President's Commission for the Study of Ethical Problems in Medicine and Biomedical and Behavioral Research, 1982, ch. 2. Many other authors--e.g., Dworkin 1988, ch. 7 & Young 2009--have also discussed the relations between autonomy and informed consent.

[7] See also Tännsjö (1999) and Jackson (2001) for similar suggestions, as well as Eyal (2014), who lays out an argument for informed consent based on building trust but attacks that argument and the idea more generally.

good course of action on the basis of her rational deliberation. So, then, medical practitioners must provide information about the probabilities of the major benefits and risks of plausible courses of action. Without information about those probabilities, patients cannot reasonably weigh the alternatives or consider tradeoffs between costs and benefits. Hence, respect for autonomy requires medical practitioners to provide information about various probabilities in order to obtain informed consent.

Proponents of alternative justifications of the requirement to obtain informed consent should also endorse the claim that patients must be provided with information about probabilities of good and bad outcomes in order to secure informed consent. First, information about the likelihood of good and bad outcomes is essential to evaluating whether *prima facie* trespasses on one's body and *prima facie* violations of one's self-ownership are, all-things-considered, justified. Second, providing the patient with statistical information about outcomes helps in preventing abusive conduct and building trust within the medical community: patients, for their part, would be 'armed' with all the relevant information, thus safeguarding them against abusive conduct (primarily by protecting patients from paternalistic practitioners who are uninformed or under-informed), and medical practitioners, for their part, would be as truthful as possible, thus promoting trust between patients and the medical community. Finally, since the risks and benefits (and their probabilities) are essential elements of any medical procedure, it is only possible for patients to waive ethical and legal requirements related to medical procedures if they are informed as to the risks and benefits. For if they are not informed as to the risks and benefits, they do not understand what requirements they are waiving. Hence, providing patients with all

the relevant information about the risks and benefits of a procedure must be part of any morally sound practice of redrawing the normative boundaries between patient and medical practitioner.

Therefore, providing patients with information about the probabilities of the risks and benefits of a procedure is required in order to secure informed consent according to the dominant justification for the informed consent requirement and according to several prominent alternative justifications for that requirement.

2. Informed Consent and Statistical Decisions

Suppose that Jack goes to his physician seeking a vasectomy. Jack's physician tells him about the risks. For example, she tells Jack that there is a small chance of about 1% that his vasectomy will fail and that his wife Jill will become pregnant after the procedure (Jamieson et al. 2004). Suppose Jack's vasectomy does fail, and Jill does become pregnant. Jack and Jill decide not to terminate the pregnancy. During labor, Jill experiences a lot of pain and asks for an epidural (central neuraxial block). Her physician tells Jill about the risks. For example, she tells Jill that there is a very small chance of about 0.0006% that an epidural will result in permanent injury (Cook et al. 2009, Table 3).[8]

Were Jack and Jill informed about their risks (assuming that an adequate *variety* of risks was discussed in ways similar to the selected examples)? We think that the answer is 'no' for three related reasons. First, as we argue in Section 2.1, the claim that some outcome has a given probability is ambiguous unless an interpretation of probability is specified, and consequently,

[8] One might think that physicians should provide only numerically precise frequency statements to patients because statements about natural frequencies are better understood (because more ecologically valid) than decimal percentages. See McDowell and Jacobs (2017) for a recent meta-analysis of work on the natural frequency facilitation effect. As with other presentation format issues, we are neutral with respect to whether probabilistic information should be presented as a natural frequency or as a decimal. We maintain that using natural frequencies does not solve all of the problems. We will return to this point later.

one cannot secure informed consent using a bare probability claim. Second, as we discuss in Section 2.2, not all probability claims are supported equally well by the evidence. In order for patients to be informed about their risks, they need to understand the quantity and quality of the evidence for the claims at stake. But even interpreted probability claims do not say anything about the quantity and quality of the evidence in their support. Hence, one cannot secure informed consent merely by stating probabilities, even if those probabilities are properly interpreted. Third, as we argue in Section 2.3, patients ultimately need to understand how probabilities may or may not be relevant to decision-making in order to be able to give genuinely informed consent. Patients need to understand the conditions under which probabilities are good guides to action. Among other things, even interpreted probability claims that are strongly supported by the evidence may be seriously misleading with respect to effective action. Hence, one cannot secure informed consent merely by stating probabilities, even if those probabilities are properly interpreted and the evidence for them is clearly described.[9]

2.1 Interpreting Probability Claims

In the long history of probability and statistics, there have been two main approaches to the interpretation of the mathematical machinery of probability: the Bayesian approach and the Frequentist approach.[10] The approach one takes to interpretation usually goes together with the approach to statistical inference one endorses and the approach to decision-making under risk

[9] Han (2013) argues that different types of probability are important to communicate to patients. He focuses primarily on the difference between what he calls "epistemic" and "aleatory" uncertainty (subjective confidence versus known risk) rather than on the difference between Bayesian and Frequentist interpretations. Like us, he concludes that greater conceptual clarity is required for adequate communication.

[10] Of course, there have been many more than these two approaches. See Hájek (2011), Hájek and Hitchcock (2016), Galavotti (2008), and Childers (2013) for introductions and discussions. Moreover, there are many ways to work out the details of each main approach. For examples of varieties of Frequentism, see Hájek (1996) and Hájek (2009). For examples of varieties of Bayesianism, see Good (1971).

that one adopts. If one is a Bayesian [Frequentist] about the interpretation of probability, then one is likely to be a Bayesian [Frequentist] about statistical inference and decision. In this section, we review the core ideas of the Bayesian and Frequentist approaches to the interpretation of probability, and we argue that reporting a bare, uninterpreted probability of success with respect to a proposed procedure or intervention is not sufficient to secure informed consent. We return to issues about statistical inference and decision in Sections 2.2 and 2.3.[11]

Traditionally, Frequentists understand probability claims to be statements about how often some event occurs within a specific collection of events, called the *reference class*. When a Frequentist says that some event type has a 1% chance of occurring, she is saying that in some specific reference class, 1 in every 100 members is a token of that event type. The reference class might be infinite, in which case the Frequentist takes the probability of an event type to be the *limiting* frequency with which it occurs. Put briefly: Probability is a measure of the (limiting) relative frequency of some type of event in a (possibly infinite) collection of events. By contrast, a Bayesian understands a probability claim to be a statement about her *credence*, by which we mean either her personal degree of belief that some proposition is true or the degree to which her evidence supports some proposition. So Bayesians but not Frequentists take probability claims to be agent-relative.

So what is it that the physician is telling Jack when she says that there is a 1% chance that his vasectomy will fail? If she is a Bayesian, she is reporting her credence that Jack's vasectomy will fail. If Jack's physician is a Frequentist, she is making a claim about the frequency with

[11] The variety of issues to consider with respect to probability, inference, and decision is staggeringly large. We cannot even list all of the most important issues here, let alone discuss them in any detail. We hope that discussing *some* of the issues will point the way for interested researchers and physicians to think more carefully and deeply about the whole range.

which vasectomies fail in some reference class.[12] Clearly these interpretations are not equivalent. The Bayesian interpretation but not the Frequentist interpretation licenses inferences regarding the physician's attitude toward a proposition. The Frequentist interpretation but not the Bayesian interpretation licenses inferences regarding repeated sampling from a specific reference class. Since the two interpretations license at least some different inferences, the question becomes pressing as to whether Jack's decision depends on the interpretation given to the chance claim.

More concretely, we can imagine a few different interpretations of the physician's claim. She might mean that around 1% of vasectomies performed by Jack's physician have failed. She might mean that 1% of vasectomies performed in Jack's country in the last 10 years have failed. Or she might mean she is personally 1% confident that Jack's vasectomy will fail. The first two interpretations are Frequentist. In the first interpretation, the reference class is patients Jack's physician has personally vasectomized. On the second, it's vasectomy patients in Jack's country in the last ten years. The third is Bayesian. On either Frequentist interpretation, Jack may think he differs in important ways from the relevant reference class. For instance, he may suspect he is younger or older, or more or less fertile than the patient population. Therefore, his personal risk could be higher or lower than the 1% figure quoted. On the Bayesian interpretation, Jack may wonder whether the physician's personal views would be widely shared amongst experts. If, for

[12] Han (2013) discusses a traditional point of dispute between Bayesians and Frequentists having to do with single-event probability. He writes (17S), "Objective probability estimates in health care are derived from and expressed in terms of the observed frequency of past outcomes in a population of individuals, and enable inferences about the frequency of expected future outcomes in a similar population. The problem, however, is that objective probability estimates are logically incoherent when applied to the prospect of a future event experienced by a single individual with only one life to live." Han treats understanding single-event probabilities as a conceptual difficulty or paradox for Frequentists, but we think that the issue is better understood as a problem for decision-making. We return to this point in Section 2.3.

instance, other experts would assess his risk at 3%, he may very well opt not to undergo the operation.

The physician has a duty to inform Jack about her interpretation of probability as part of obtaining his consent, since Jack's decision plausibly depends on the interpretation. Knowing only a bare probability value does not provide a complete enough basis for rational decision-making, including waiving normative requirements. If so, then knowing only a bare probability value is not sufficient to secure informed consent.

Physicians might be tempted to say that there is no ambiguity in their probability claims because *obviously* they are reporting actual frequencies. But we are not so sure. Suppose a physician has access to several related, but only *partially informative* studies. For example, a physician might be tasked with treating an elderly Canadian of Inuit descent but only have evidence from a study on African Americans living in the United States and a study on elderly people living in Australia. When a physician has access to one or more studies involving participants who are similar to the patient in many respects but also different from the patient in potentially important ways, frequency information does not directly bear on the question at issue. In such cases, the physician's chance claim is best interpreted as a report of credence with respect to the proposition that the intervention will be successful. Similarly, suppose a physician is trying to communicate the risks with respect to a procedure that has never been tried or that has been tried only a very small number of times. In such cases, if the physician says that the procedure has some probability of success, she isn't making any claim about *actual* frequencies. She might be making a claim about *hypothetical* frequencies, but we think it is more likely that she is simply expressing her credence that the procedure will be successful. Moreover, even if

physicians *intend* to report actual frequencies in all cases where they report probabilities, they need to present information in a way that is intelligible to their patients. If physicians report probabilities that they interpret as actual frequencies but patients understand those reports in terms of credences, then physician reports will not secure informed consent.[13]

2.2 On the Evidence for Probability Claims

In the previous section, we reviewed the core ideas of the Bayesian and Frequentist approaches to the interpretation of probability, and we argued that reporting a bare, uninterpreted probability of success for a proposed medical procedure or intervention is not sufficient to secure informed consent. In this section, we argue that even an *interpreted* probability claim is not sufficient to secure informed consent. In order for patients to be informed about their risks, they need to understand the quantity and quality of the evidence for the probabilities that their physicians report.[14] Cook et al. (2009, 180) suggest a particularly strong version of such a requirement. Lamenting the lack of reliable estimates for the frequency of injuries from neuraxial block, they say that physicians need to know and accurately report the frequency with which complications arise in order to obtain genuinely informed consent:

> Knowledge of the incidence of such complications [as permanent injury from neuraxial block] should be an essential component of the clinical decision-making and consent processes, but there are few good data which can be quoted to support such discussions, leaving both patient and clinician in a quandary. Figures (ranging from 1:1000 to 1:100

[13] One might think of the issue here as a consequence of the fact that consent is not transitive, for which see O'Neill (2003). A patient might assent to undergoing a procedure that has a 70% probability of success but not a procedure that succeeds 70 times out of 100.

[14] What we have in mind here is similar to what is sometimes called "ambiguity" (see Han 2013). However, we think there is an important distinction to be made between evaluating the quantity and quality of evidence for a probability claim and the complete uncertainty about probability (what we would count as genuine ambiguity) that figures in Ellsburg's paradox and related puzzles in decision theory. While Han (2013) worries that ambiguity aversion will lead patients to avoid decision making if they are presented with higher-order probabilities, the studies discussed and presented in Clarke et al. (2015) suggest that in cases where the quantity and quality of evidence is good, "including weight of evidence content … attenuates perceived information uncertainty" (1302).

000) are quoted, but their doubtful validity questions the ability to obtain genuinely informed consent from patients offered these procedures.

Cook et al. do not provide an argument for the claim that physicians need to know the frequency with which complications arise in order to obtain genuinely informed consent from their patients, but we take the following to be a charitable attempt to lay out an argument on their behalf.

[C1] If physicians do not know the frequency with which complications occur for a medical intervention, then physicians cannot non-accidentally report a true and reliable estimate of the frequency with which complications occur for that intervention.

[C2] Physicians can satisfy their obligation to secure informed consent with regard to a medical intervention only if they can non-accidentally report a reliable estimate of the frequency with which complications occur for that intervention.

[C3] If physicians do not know the frequency with which complications occur for a medical intervention, then they cannot satisfy their obligation to secure informed consent with regard to that intervention.

The standard advanced by Cook et al. is very strict. The physician has to *know* the frequency of complications. Cook et al. don't specify in any detail how much evidence is required to know such a thing. But their procedure for obtaining an estimate gives some indication. They attempted to "identify both numerator (number of major complications) and denominator (number of CNB) information for a 12 month period by a review across the breadth of anaesthetic and pain management practice in the UK National Health Service (NHS)" (180). Something that Cook et al. do not address is whether and to what extent physicians need to *communicate* the quantity and quality of their evidence. We think that communicating the quantity and quality of the evidence is indispensable for securing informed consent. Let's consider an example.

Jill's physician tells her that there is a 0.0006% probability of permanent injury from an epidural, and she explains that her probability claim is expressing the frequency with which permanent injuries occur in such procedures. That strikes Jack and Jill as an acceptably low chance of injury. But Jack wants to know more. After all, Jack thought that the probability of his vasectomy working was high, but Jill got pregnant anyway. Once bitten, twice shy, Jack asks for more information about the physician's claim that there is a 0.0006% probability of permanent injury from an epidural. The physician says, "Look, Jack, there was a large study of complications from epidurals; and the study was explicitly trying to determine the frequency of permanent injury in these procedures." The physician then refers Jack to Cook et al. (2009) who report that there were 161,550 obstetric epidural procedures performed in the UK during the year covered by their study with only one patient suffering permanent injury.

How good is the evidence that Jill has a very small chance of suffering a permanent injury from her epidural? Clearly it is much stronger than if Cook et al. had sampled 100 patients. A little bad luck in a sample of that size would produce an estimate more than 1500 times too large (assuming that Cook et al.'s estimate is correct). It's not unusual for medical studies to vary widely in sample size. For example, the estimate of vasectomy failure from Jamieson et al. (2004) was based on 540 women. But the quantity and quality of the evidence for a probability claim is relevant to whether and how the claim should guide action. Suppose two researchers collect data on the success of a procedure. The first researcher studies 100 people and observes 1 failure. The other researcher studies 10,000 people and observes 100 failures. The point estimates of the probability of success will be 0.99 in both cases. Assuming that the sampling procedures are the same, the estimate is much more secure for the researcher with the

larger sample.[15] If so, then physicians need to communicate *something* about the quantity and quality of the evidence for the probabilities they report in order to secure informed consent. But exactly *what* should physicians communicate about the quantity and quality of their evidence, and exactly *how* should they do so?

Physicians could adopt the usual Frequentist strategy and describe the reliability of the statistical method used in producing the estimate. This would fit well with the fact that so much of the medical literature uses Frequentist statistical tools. For example, Jill's physician might point to the 95% confidence interval of (0, 3.4) that Cook et al. report for their point estimate of 0.6 permanent injury events per every 100,000 cases.[16] "You see, Jack," she might say, "if lots of researchers gathered data about the frequency of permanent injuries from neuraxial block, then about 95 out of every 100 intervals constructed according to the method used by Cook et al. would have the true frequency in its range." But Frequentist confidence intervals are difficult to understand, even when they have been carefully explained in detail. Hoekstra et al. (2014) provide empirical evidence that many researchers in psychology—a discipline that makes extensive use of Frequentist statistical tools—do not have a solid understanding of confidence intervals.[17]

---

[15] Peirce (1878) used the label "weight of evidence" to refer to what we would call the "quantity" of the evidence. For further discussion, see Keynes (1921, Ch. 6), Good (1985), and Bradley (2016, Ch. 14). For discussion of relatively recent use of the phrase "weight of evidence" in biomedical science, see Weed (2005).

[16] [Redacted] (personal communication) reminded us that an interval only expresses weight of evidence relative to some assumed level of certainty or confidence. Just how much certainty or confidence one wants when making a decision will depend on what is at stake. We demand more certainty or confidence for decisions about what procedure to use for heart surgery than for decisions about what type of cast to use for a broken leg. This point applies to both Frequentist and Bayesian approaches to interval estimation.

[17] Hoekstra et al. modeled their study on a famous study by Gigerenzer (2004), which provided empirical evidence that researchers frequently misinterpret p-values, another mainstay of Frequentist statistics. In their study, Hoekstra et al. told a story about a professor who reports a 95% confidence interval of (0.1, 0.4) for a mean value being estimated. They then asked 118 researchers to say whether each of six statements was true or false. The number of researchers reporting the wrong answer ranged from 45 to 102. For example, 70 out of 118 endorsed the false claim that there is a 95% probability that the true mean lies between 0.1 and 0.4, and 68 out of 118 endorsed the false claim that if we were to repeat the experiment over and over, then 95% of the time the true mean falls between 0.1

If people who are familiar with statistics—people who learn statistics as part of their professional training and who use statistics every day in their research work—misunderstand confidence intervals at such high rates, then physicians cannot expect typical patients to find confidence intervals intuitive. One might be tempted to think that physicians should avoid inferential statistics altogether and report only raw, observed frequencies. Patients would then have to draw their own conclusions. However, physicians would then have to be sure that they were providing all of the information *needed* for patients to draw their own conclusions. For one thing, physicians would need to report facts about experimental designs, such as how the data were sampled, since the sampling procedure sometimes matters for the inference drawn.[18]

It may seem, then, that doctors should abandon frequencies and instead report their (Bayesian) credences on the assumption that such reports are easier to understand. However, Bayesian credences are not without problems of their own. Here, we want to illustrate two problems having to do with communicating the evidence for a probability claim under a Bayesian interpretation. The first problem is that two (or more) very different bodies of evidence may produce the same credence. Jill wants to know the probability that she will suffer a permanent injury. If her doctor were a Bayesian, she would calculate the probability of permanent injury given her evidence. Suppose Jill's doctor started her career thinking that all of the possible values for the probability of success were equally likely. That is, she thought it was just as likely that permanent injuries were very probable as that they were extremely unlikely and so on. She then observed some number of trials, all of which she regarded as independent. If

and 0.4. The first of these mistakes a Frequentist confidence interval for a Bayesian credible interval. The second treats the confidence interval as fixed and the parameter being estimated as variable; whereas, the Frequentist thinks of the interval as variable and the parameter as fixed.

[18] Even if one is a Bayesian, the sampling procedure may matter in special cases. See Steel (2003).

Jill's doctor has seen one permanent injury in the course of 98 trials, she calculates the probability that Jill will suffer a permanent injury as 2%.[19] Jill's doctor would *also* have reported a 2% chance of permanent injury if she had observed no permanent injuries in 48 trials or if she had observed 3 permanent injuries in 198 trials. But those bodies of evidence lead to very different credence *distributions*, and hence, they are not all equally informative. In order to communicate the quantity and quality of her evidence, Jill's physician might specify a range—a *credible interval*—such that she has a high credence that the frequency of permanent injuries is in the interval. For example, she might be willing to bet at 20 to 1 odds that the number of permanent injuries per 100,000 is in the interval (0, 4).[20] That would be an expression of her credence regarding the frequency, and it would represent her assessment of the quantity and quality of the evidence. But it would go well beyond reporting an interpreted probability value.

More commonly, a doctor won't have such readily articulable evidence, which raises problems of transparency and expert disagreement. To illustrate, suppose that Richard is a morbidly obese 66 year old male considering whether to continue to rely on a urinary catheter or undergo urethral reconstructive surgery because of stricture. He has a long history of urinary tract infections that are resistant to antibiotics and is also on immuno-suppresants because he has received a kidney transplant. He has had Crohn's disease since his twenties and had large parts of his intestines removed decades ago. Richard asks his doctor how likely it is that the operation

[19] In the circumstances described, she applies Laplace's Rule of Succession and says that the probability of injury on the next procedure is equal to $(m + 1) / (n + 2)$, where m is the number of injuries observed out of n procedures. See Zabell (2005), especially Chapter 2, for more on the Rule of Succession.

[20] Traditionally, Bayesian credences have been interpreted as betting odds. If you are 70% sure that an event will occur, it means (roughly) you personally are willing to pay up to 70 cents for a bet that returns $1 if the event occurs and nothing otherwise. Although such an interpretation is illustrative, it is not without problems. In particular, it seems distasteful (at the least) to bet on whether a patient's operation will be successful. Nonetheless, we think the betting odds framework helps highlight the difference between the Frequentist and Bayesian approaches.

will be a success. Because she knows that he has complications that make him importantly different from the average patient undergoing this operation, she does not report historical frequencies of success. Instead, she takes his history into account and reports her personal confidence level of 70% that the operation will be successful. She might even include a 99% credible interval in her report in order to more accurately communicate the quantity and quality of her evidence. But despite the fact that Richard's doctor is an expert with extensive clinical experience, her figure of 70% was (likely) not arrived at through any careful or transparent means. And importantly, her initial credence is not open to inspection or criticism. If Richard were to ask a second physician who is equally qualified, she may very well report a different number. More generally, if the experts' credences are highly sensitive to their priors, then there will be a lot of possible variation between the credences of different experts. In such cases, experts have some duty to inform patients that other experts may have or likely *will* have different credences. Again, the result is that even reporting an interpreted probability is not sufficient to secure informed consent.[21]

We have so far focused on issues of data *quantity*, but quality matters, too. A study having 100 participants selected at random from a population to which a target patient belongs is much better than a study having 100 participants selected systematically or selected from a population to which the target patient does not belong. Similarly, data collected in a controlled trial will be more informative than observational data. In order to secure informed consent, physicians should report information about the quality of their evidence, as well as its quantity.

[21] Professional or expert judgment about risk is an important issue in areas outside biomedical ethics, as well. See Murphy and Gardoni (2008) for some discussion of expert judgment in engineering ethics. In certain cases, physicians may well come close to consensus even if they have trouble articulating precisely what their evidential basis is. However, since physicians disagree relatively often, such cases are far from universal.

2.3 Decisions, Decisions

In the previous two sections, we argued that reporting bare, uninterpreted probability values is not sufficient to secure informed consent, and we argued that in addition to including an interpretation, patients need to understand the quantity and quality of the evidence for the probabilities that physicians report in order to give genuinely informed consent. Now suppose a physician reports an interpreted probability that some procedure will be successful, and suppose she also adequately conveys the quantity and quality of the evidence for her probability claim. Will her reports be sufficient to secure informed consent? In this section, we argue that they will not be. A report may be misleading because the probability value reported is not action-guiding.

Suppose that Jack develops kidney stones. He goes to see a local nephrologist, who describes two possible surgical procedures. One involves open surgery, and the other is minimally invasive. The nephrologist tells his patients that open surgery has a success rate of 72% and that the minimally invasive surgery has a success rate of 83%. He tells his patients that these probabilities should be understood in terms of frequencies, and he tells his patients that his probability claims are based on a well-designed study involving more than 400 subjects. Everything he says is true, and as a result of his report, essentially all of his patients choose the minimally invasive option. Jack is not unusual in this regard, and he chooses the minimally invasive surgery as well. But unknown to Jack, the nephrologist is not an honest man. The nephrologist knows that this is a case of Simpson's Paradox and that among patients with small kidney stones (< 2cm), open surgery has a success rate of 92% compared to 87% for the minimally invasive option, and among patients with large kidney stones (≥ 2cm), open surgery

has a success rate of 71% compared to 69% for the minimally invasive option.[22] However, he only ever reports aggregated probabilities because he makes more money performing minimally invasive surgeries, and he knows, let us say, that if he were to report the success rates separated out by size of stone, his patients would be much more likely to choose open surgery.[23]

We think it is clear that the nephrologist acted unethically and that Jack did not give genuinely informed consent for his surgery. The physician knows (and withholds) something that is relevant to Jack's decision. Specifically, it is plausible that conditioning on the size of the kidney stones discloses something important about the causal structure, and the physician knows (or suspects) that this is the case.[24] In the case of the dishonest doctor (and in many other cases where we need to choose an *action*), probabilities are relevant to decisions only insofar as they track the causal structure, and hence, the aggregated probabilities that the doctor reports are misleading.[25]

Our story about Jack and the unscrupulous nephrologist shows that even reporting a properly interpreted probability value together with an accurate statement about the quantity and quality of the evidence for that probability value is not sufficient to secure informed consent. But our tale is one of obvious abuse, and the problematic feature of the account is pretty clear. Jack's

[22] Simpson's Paradox occurs when an association between two variables disappears or reverses conditional on a third variable. The threat of Simpson's Paradox is non-trivial. Hanley and Thériault (2000) reported an example of Simpson's Paradox in meta-analyses of randomized controlled trials. Nissen and Wolski (2007) found a significant increase in the risk of myocardial infarctions in groups taking Rosiglitazone over the control groups over a number of individual studies. However, when data from all the studies were pooled, the Rosiglitazone group had a slightly lower rate of MIs compared to the control.

[23] The numbers for our toy example are based on a famous real-life study by Charig et al. (1986) that compared different methods for treating kidney stones. The unscrupulousness is novel to our story.

[24] For our purposes, a "causal structure" is a pattern of causal relations that hold with respect to some domain. Minimally, we take a causal relation to tell us how an outcomes depends on changes in actions we could take. We want to know, for example, what would happen to success rate if we were to choose open surgery rather than a less invasive option. See Cartwright (1979) and Woodward (2003) for introductions to some of the issues here.

[25] In the language of Pearl's (2000) do calculus, a report of the value of $Pr(Y=y \mid X=x)$ is action-guiding only if it is a guide to the value of $Pr(Y=y \mid do(X=x))$.

physician knew something that was relevant to Jack's decision but withheld that information. Failing to provide information due to negligent ignorance would also block genuinely informed consent. But what if a physician neither withholds information intentionally nor fails to provide it due to negligence but simply doesn't know the relevant information? Physicians surely do not always need to know the causal structure of a given case in order to properly inform their patients. However, physicians should have (and should express that they have) no reason to think that the probabilities they report are not action-guiding. That is, they should think the probabilities they provide are the best information available for rational decision-making. Moreover, physicians should explain how and why probability values sometimes are and sometimes are not decision-guiding. Patients may not be aware of Simpson's Paradox or related issues. If a patient knows only that success occurs with a specific probability value for some sample population, she may not know that the probability of success could be very different for an alternative subgroup to which she belongs and that subgroup membership tracks the decision-relevant causal facts.

A similar problem arises when either the sampled population is heterogeneous or idiosyncratic features of the patient are potentially highly relevant. Suppose Jack and Jill were deciding between no sterilization for either of them, a vasectomy for Jack, or a tubal sterilization for Jill. Suppose Jack would have another child if Jill wanted, but Jill is fairly sure she doesn't want one. Nonetheless, she worries that she'll regret sterilization. Jack and Jill ask their doctor for advice. The doctor points out that Jamieson et al. (2002) found that after five years, 6.1% of women studied whose husband had a vasectomy said they regretted the procedure, whereas 7% of women expressed regret five years after tubal ligation.

Although these frequencies are potentially useful, they can also be highly misleading. Jack and Jill are, like any couple, idiosyncratic. They have features that make them importantly different from any member of the sampled population. For instance, suppose Jill is very decisive and rarely changes her mind but also has family who would be upset if they learned Jill decided to have any sterilization procedure done. She and Jack also have four children already---more than the average American family---and additional children would add a significant financial burden after Jack was recently laid off. Mere frequency data should be used with extreme caution in this case.

The upshot is that reporting an interpreted probability that some procedure will be successful is not sufficient to secure informed consent even if the physician adequately conveys the quantity and quality of the evidence supporting her probability claim. Not all probabilities are action-guiding, and we can have better or worse reasons to think that a probability value is action-guiding. Hence, it's important to convey information about the decision-relevance of probability values to patients in order to obtain genuinely informed consent.

3. Conclusion: A Dilemma for the Informed Consent Theorist

The literature on informed consent in medicine recognizes that information about the probabilities of success and failure associated with possible treatments must be communicated to patients in order to obtain genuinely informed consent. In this paper, we have argued that more is required. First, patients need to understand the way in which probability claims are being interpreted. Second, patients need to understand the quantity and quality of the evidence supporting probability claims. Third, patients need to understand how the reported probabilities

may or may not be relevant to decision-making. All three of these elements are missing when physicians report bare, uninterpreted probabilities.

The upshot of our discussion is that the informed consent theorist faces a dilemma. According to the first horn of the dilemma, the informed consent theorist ought to bite the bullet and agree that patients should not be informed of probabilities: it is neither feasible for physicians to explain nor for patients to understand the way in which probability claims are interpreted, the quantity and quality of the evidence supporting the claims, and the relevance of the reported probabilities to their decision-making. In other words, since it is too difficult to educate physicians and patients about the various interpretations of probability, issues surrounding statistical inference, and the relations between probabilities and decision-making, there is simply no point in presenting probability-related information to patients. If this horn of the dilemma is adopted, then the literature on informed consent would have to be revised so as to remove talk of the necessity, or even the importance, of reporting probability claims to patients.

According to the second horn of the dilemma, the informed consent theorist might insist that reporting probabilities is necessary for obtaining genuinely informed consent, but argue that physicians can provide the required information about how the physician's probability claims are to be understood, about the quantity and quality of the evidence for those claims, and about the relevance of those claims to decision-making. However, preparing physicians to do so is a difficult task that will require resources and planning. Physicians would need to be trained much more substantially in statistics and the philosophical foundations of statistics. Since even active researchers who use statistics daily often misunderstand basic statistical tools and techniques, such as confidence intervals and the like, clinicians cannot currently be expected to understand

such concepts fully either. Moreover, physicians would need to be trained on the relationship between probability and statistics on the one hand and decision-making on the other. Additional courses in statistics, in causal reasoning, in decision theory, and in the philosophy of probability and statistics, along with courses or workshops on conveying statistical information to patients would have to be added to the curricula of medical schools (and other medical programs). Furthermore, third-party patient activist decision theorists might be needed to help patients understand how probabilities relate to their decisions. Such theorists could be members of broader patient activist groups and would help patients understand how the probabilities presented to them pertain to the decision at hand. All of this would require substantial resources, of course, but it is a way in which the informed consent theorist could retain a requirement to report probabilities as part of her theory.

Now, it might be the case that the informed consent theorist would attempt to reject the dilemma by adopting a "something-is-better-than-nothing approach," according to which reporting probability claims is better than not reporting them, *even if* they remain uninterpreted and/or the quantity and quality of the evidence supporting the probability claims is not provided and/or the relevance of these claims to the patient's decision-making is not explained. On this scenario, the informed consent theorist might argue that even if it is too difficult to educate physicians and patients about all the aspects of reporting and understanding probabilities, providing the patient with *some* information about probabilities is better than not providing them with any such information.

However, given the argument of this paper, it is far from clear that providing patients with anything less than fully-interpreted probability claims along with the relevant evidence

supporting these claims and an explanation of how the claims pertain to their decision making, satisfies the justification provided for informed consent in the literature. Consider the autonomy justification for informed consent: given the issues we have raised in the paper,if the physician provides an uninterpreted probability, she will not be providing information that is helpful to the patient in choosing a good course of action: information about uninterpreted probabilities does not help patients in reasonably weighing alternatives or considering tradeoffs between costs and benefits. And this means that this type of information does not really enhance patients' autonomy. Moreover, recall that proponents of alternative justifications of the requirement to obtain informed consent also had reason to endorse the claim that patients must be provided with information about probabilities of good and bad outcomes in order to secure informed consent. However, if the information that is provided to patients is unhelpful to their decision process, it is far from obvious how it might help patients assess whether *prima facie* trespasses on their bodies and *prima facie* violations of their self-ownership are justified or whether they should waive ethical and legal requirements related to medical procedures. Moreover, without fully understanding the probability claims, such information would do little to prevent abusive conduct or to build trust within the medical community. Therefore, the "something-is-better-than-nothing approach" is untenable. Indeed, this suggestion collapses into the first horn in the sense that the informed consent theorist will need to give up on the requirement of reporting probability claims.

Informed consent researchers face a dilemma that is not easily avoidable: either the requirement to report the uncertainty of medical interventions must be substantially weakened or the medical community will need to invest much more heavily in training and staffing for

satisfying the demands of informed consent with respect to uncertainty in medical interventions. We are not sure which horn to take. What we are sure of is that researchers interested in medical ethics -- and especially those interested in informed consent -- need to think more carefully about probability.

References

Archard, D. (2008). Informed Consent: Autonomy and Self-Ownership. *Journal of Applied Philosophy*, 25(1): 19–34.

Beauchamp, T. L. (2010). Autonomy and consent. In *The Ethics of Consent*. F. G. Miller & A. Wertheimer (Eds.). New York: Oxford University Press.

Beauchamp, T. L., & Childress, J. F. (2008). *Principles of Biomedical Ethics* (6th edition), Oxford: Oxford University Press.

Benjamin, D. J., Berger, J. O., Johannesson, M., Nosek, B. A., Wagenmakers, E. J., Berk, R., ... & Cesarini, D. (2018). Redefine statistical significance. *Nature Human Behaviour*, 2(1): 6-10.

Berg, J. W., Appelbaum, P. S., Lidz, C. W., & Parker, L. S. (2001). *Informed Consent: Legal Theory and Clinical Practice* (2nd Edition). New York: Oxford University Press.

Bradley, R. (2016). *Decision Theory with a Human Face*. Cambridge University Press.

Brock, D. W. (1987). Informed Consent. In *Health Care Ethics: An Introduction*. D. VanDeVeer & T. Regan (Eds.). Philadelphia: Temple University Press.

Cartwright, N. (1979). Causal laws and effective strategies. *Nous*, 13(4): 419-437.

Childers, T. (2013). *Philosophy and probability*. Oxford University Press.

Christopher, M. M., Hotz, C. S., Shelly, S. M., & Pion, P. D. (2010). Interpretation by clinicians of probability expressions in cytology reports and effect on clinical decision‐making. *Journal of veterinary internal medicine*, 24(3): 496-503.

Clarke, C., McKeever, B., Holton, A., & Dixon, G. (2015). The Influence of Weight-of-Evidence Messages on (Vaccine) Attitudes: A Sequential Mediation Model. *Journal of Health Communication*, 20(11): 1302-1309.

Collins, P. J., & Hahn, U. (2018). Communicating and reasoning with verbal probability expressions. *Psychology of Learning and Motivation*, 69: 67-105.

Cook, T. M., Counsell, D., & Wildsmith, J. A. W. (2009). Major complications of central neuraxial block: report on the Third National Audit Project of the Royal College of Anaesthetists. B*ritish Journal of Anaesthesia*, 102(2): 179-190.

Dieckmann, N. F., Mauro, R., & Slovic, P. (2010). The effects of presenting imprecise probabilities in intelligence forecasts. *Risk Analysis: An International Journal*, 30(6): 987-1001.

Dworkin, G. (1988). *The Theory and Practice of Autonomy*. Cambridge & New York: Cambridge University Press.

Emanuel, E. J., & Emanuel, L. L. (1992). Four Models of the Physician-Patient Relationship. *The Journal of the American Medical Association*, 267(16): 2221-2226.

Eyal, N. (2014). Using informed consent to save trust. J*ournal of Medical Ethics*, 40(7): 437-444.

Faden, R. R. & Beauchamp, T. L. (1986). *A History and Theory of Informed Consent*. New York: Oxford University Press.

Galavotti, (2008). Probability. In *The Routledge Companion to Philosophy of Science*. M. Curd & S. Psillos (Eds.). Abingdon, Oxon: Routledge.

Gigerenzer, G. (2004). Mindless statistics. *The Journal of Socio-Economics*, 33(5): 587-606.

Good, I. J. (1971). 46656 varieties of Bayesians. *American Statistician*, 25(5), 62-63.

Good, I. J. (1985). Weight of Evidence: A Brief Survey. *Bayesian Statistics*, 2: 249-269.

Hájek, A. (1996). "Mises Redux" -- Redux: Fifteen Arguments against Finite Frequentism. *Erkenntnis*, 45(2-3): 209-227.

Hájek, A. (2009). Fifteen Arguments Against Hypothetical Frequentism. *Erkenntnis*, 70(2): 211-235.

Hájek, A. (2012). Interpretations of Probability. T*he Stanford Encyclopedia of Philosophy*, E. N. Zalta (Ed.). https://plato.stanford.edu/archives/win2012/entries/probability- interpret.

Hájek, A., & Hitchcock, C. (2016). Probability for Everyone—Even Philosophers. In *The Oxford Handbook of Probability and Philosophy*. A. Hájek & C. Hitchcock (Eds). Oxford: Oxford University Press.

Hamann, J., Mendel, R., Bühner, M., Leucht, S., & Kissling, W. (2011). Drowning in numbers–what psychiatrists mean when talking to patients about probabilities of risks and benefits of. medication. *European Psychiatry*, 26(2): 130–131.

Han, P. K. J. (2013). Conceptual, Methodological, and Ethical Problems in Communicating Uncertainty in Clinical Evidence. *Medical Care Research and Review, Supplement*, 70(1): 14S-36S.

Han, P. K. J., Joekes, K., Elwyn, G., Mazor, K. M., Thomson, R., Sedgwick, P., Ibison, J., & Wong, J. B. (2014). Development and evaluation of a risk communication curriculum for medical students. *Patient Education and Counseling*, 94(1): 43–49.

Hanauer, D. A., Liu, Y., Mei, Q., Manion, F. J., Balis, U. J., & Zheng, K. (2012). Hedging their mets: the use of uncertainty terms in clinical documents and its potential implications when sharing the documents with patients. *AMIA Annual Symposium Proceedings*, 2012: 321-330.

Hanley, J. A., & Thériault, G. (2000). Simpson's paradox in meta-analysis. *Epidemiology*, 11(5): 613.

Hoekstra, R., Morey, R. D., Rouder, J. N., & Wagenmakers, E. J. (2014). Robust misinterpretation of confidence intervals. *Psychonomic Bulletin & Review*, 21(5): 1157-1164.

Jackson J. (2001). *Truth, trust and medicine*. London and New York: Routledge.

Jamieson, D. J., Kaufman, S. C., Costello, C., Hillis, S. D., Marchbanks, P. A., Peterson, H. B., & US Collaborative Review of Sterilization Working Group. (2002). A comparison of women's regret after vasectomy versus tubal sterilization. *Obstetrics & Gynecology*, 99(6): 1073-1079.

Jamieson, D. J., Costello, C., Trussell, J., Hillis, S. D., Marchbanks, P. A., and Peterson, H. B. (2004). "The Risk of Pregnancy After Vasectomy." Obstetrics & Gynecology 103(5), 848-850.

Jenkins, S. C., Harris, A. J., & Lark, R. M. (2018a). Understanding 'unlikely (20% likelihood)'or '20% likelihood (unlikely)'outcomes: The robustness of the extremity effect. *Journal of Behavioral Decision Making*, 31(4): 572-586.

Jenkins, S. C., Harris, A. J., & Lark, R. M. (2018b). When unlikely outcomes occur: the role of communication format in maintaining communicator credibility. *Journal of Risk Research*, 1-18. doi: 10.1080/13669877.2018.1440415

Karelitz, T. M., & Budescu, D. V. (2004). You say "probable" and I say "likely": improving interpersonal communication with verbal probability phrases. *Journal of Experimental Psychology: Applied*, 10(1): 25-41.

Katz, J. (1984). *The Silent World of Doctor and Patient*. New York: The Free Press.

Keynes, J. (1921). *A Treatise on Probability*. London: Macmillan and Company.

Kong, A., Barnett, G. O., Mosteller, F., & Youtz, C. (1986). How Medical Professionals Evaluate Expressions of Probability. *New England Journal of Medicine*, 315(12): 740-744.

Lakens, D., Adolfi, F. G., Albers, C. J., Anvari, F., Apps, M. A., Argamon, S. E., ... & Buchanan, E. M. (2018). Justify your alpha. *Nature Human Behaviour*, 2(3): 168-171.

Levine, R. J. (1988). *Ethics and Regulation of Clinical Research* (2nd edition). New Haven, CT: Yale University Press.

Lipkus, I. M. (2007). Numeric, verbal, and visual formats of conveying health risks: suggested best practices and future recommendations. *Medical Decision Making*, 27(5): 696-713.

Longman, T., Turner, R. M., King, M., & McCaffery, K. J. (2012). The effects of communicating uncertainty in quantitative health risk estimates. *Patient Education and Counseling*, 89(2): 252-259.

Manson, N., and O'Neill, O. (2007). *Rethinking Informed Consent in Bioethics*. Cambridge: Cambridge University Press.

Mapes, R.E.A. (1979). Verbal and Numerical Estimates of Probability in Therapeutic Contexts. *Social Science and Medicine*, 13(A): 277-282.

Martin, S. S, Sperling, L. S., Blaha, M. J., Wilson, P. W. F., Gluckman, T. J., Blumenthal, R. S., & Stone, N. J. Clinician-Patient Risk Discussion for Atherosclerotic Cardiovascular Disease Prevention. *Journal of the American College of Cardiology*, 65(13): 1361-1368.

Mazur, D. J., & Jon F. Men, J. F. (1994). Patients' Interpretations of Verbal Expressions of Probability: Implications for Securing Informed Consent to Medical Interventions. *Behavioral Sciences and the Law*, 12(4): 417-426.

McDowell, M. & Jacobs, P. (2017). Meta-Analysis of the Effect of Natural Frequencies on Bayesian Reasoning. *Psychological Bulletin*, 143(12): 1273-1312.

Merz, J. F., Druzdzel, M. J., & Mazur, D. J. (1991). Verbal Expressions of Probability in Informed Consent Litigation. *Medical Decision Making*, 11(4): 273-281.

Millum, J., and Bromwich, D. (2013). Disclosure and Consent to Medical Research Participation. *Journal of Moral Philosophy*, 10(4): 195–219.

Millum, J., and Bromwich, D. (2018). Understanding, Communication, and Consent. *Ergo* 5(2): 45-68.

Murphy, C., & Gardoni, P. (2008). The acceptability and the tolerability of societal risks: A capabilities-based approach. *Science and Engineering Ethics*, 14(1): 77-92.

Nakao, M.A. & Axelrod, S. (1983). Numbers are Better than Words: Verbal Specifications of Frequency Have No Place in Medicine. *The American Journal of Medicine*, 74(6): 1061-1065.

National Commission for the Protection of Human Subjects of Biomedical and Behavioral Research. (1979). The Belmont Report. *Federal Register*, 44(76): 23192–23197.

Nelson, W., Reyna, V. F., Fagerlin, A., Lipkus, I., & Peters, E. (2008). Clinical Implications of Numeracy: Theory and Practice. *Annals of Behavioral Medicine*, 35(3): 261–274.

Neyman, J. (1957). "Inductive Behavior" as a Basic Concept of Philosophy of Science. *Revue de l'Institut International de Statistique / Review of the International Statistical Institute*, 25(1/3): 7-22.

Nissen, S. E., & Wolski, K. (2007). Effect of rosiglitazone on the risk of myocardial infarction and death from cardiovascular causes. *New England Journal of Medicine*, 356(24): 2457-2471.

O'Neill O. (2002). *Autonomy and trust in bioethics*. Cambridge: Cambridge University Press.

O'Neill, O. (2003). Some limits of informed consent. *Journal of Medical Ethics*, 29(1), 4-7.

Pearl, J. (2000). *Causality*. Cambridge University Press.

Peirce, C.S. (1878). The probability of induction. *Popular Science Monthly*, 12: 705-718.

President's Commission for the Study of Ethical Problems in Medicine and Biomedical and Behavioral Research (1982). *Making Health Care Decisions: A Report on the Ethical and Legal Implications of Informed Consent in the Patient-Practitioner Relationship*. Washington, D.C.: The Commission.

Rücker, G., & Schumacher, M. (2008). Simpson's paradox visualized: the example of the rosiglitazone meta-analysis. BMC medical research methodology, 8(1): 34.

Savulescu, J. (1995). Rational Non-Interventional Paternalism: Why Doctors Ought to Make Judgments of What is Best for Their Patients. *Journal of Medical Ethics*, 21(6): 327-331.

Sladakovic, J., Jansen, J., Hersch, J., Turner, R., & McCaffery, K. (2016). The differential effects of presenting uncertainty around benefits and harms on treatment decision making. *Patient Education and Counseling*, 99(6): 974–980.

Spiegelhalter, D., Pearson, M., & Short, I. (2011). Visualizing uncertainty about the future. *Science*, 333(6048): 1393-1400.

Spiegelhalter, D. (2017). Risk and uncertainty communication. *Annual Review of Statistics and Its Application*, 4: 31-60.

Steel, D. (2003). A Bayesian way to make stopping rules matter. *Erkenntnis*, 58(2): 213-227.

Tait, A. R., Voepel-Lewis, T., Zikmund-Fisher, B. J., & Fagerlin, A. (2010). Presenting research risks and benefits to parents: does format matter?. *Anesthesia and analgesia*, 111(3): 718-723.

Tännsjö T. (1999). *Coercive care: the ethics of choice in health and medicine*. London: Routledge.

Theil, M. (2002). The role of translations of verbal into numerical probability expressions in risk management: a meta-analysis. *Journal of Risk Research*, 5(2): 177–186.

Vahabi, M. (2010). Verbal versus numerical probabilities: Does format presentation of probabilistic information regarding breast cancer screening affect women's comprehension? *Health Education Journal*, 69(2): 150–163.

Weed, D. (2005). Weight of Evidence: A Review of Concept and Methods. *Risk Analysis*, 25(6): 1545-1557.

Woodward, J. (2003). *Making Things Happen*. Oxford: Oxford University Press.

Young, R. (2009). Informed Consent and Patient Autonomy. In *A Companion to Bioethics* (2nd Edition). H. Kuhse & P. Singer (Eds). Malden MA & Oxford: Blackwell Publishing Ltd.

Zabell, S. L. (2005). *Symmetry and its discontents: essays on the history of inductive probability*. Cambridge University Press.

Zipkin, D. A., Umscheid, C. A., Keating, N. L., Allen, E., Aung, K., Beyth, R., ... & Schardt, C. (2014). Evidence-based risk communication: a systematic review. *Annals of internal medicine*, 161(4): 270-280.