An Analysis of the Epidemiology and Related Evidence on the Health Effects of Wind Turbines on Local Residents

prepared at the request of Brown County Citizens for Responsible Wind Energy in connection with Public Service Commission of Wisconsin docket no. 1-AC-231, Wind Siting Rules

Carl V. Phillips, MPP PhD
epiphi Consulting Group
cvphilo@gmail.com

3 July 2010

Executive Summary

A summary of the main conclusions of my expert opinion, based on my knowledge of epidemiology and scientific methods, and my reading of the available studies and reports, is as follows:

- There is ample scientific evidence to conclude that wind turbines cause serious health problems for some people living nearby. Some of the most compelling evidence in support of this has been somewhat overlooked in previous analyses, including that the existing evidence fits what is known as the case-crossover study design, one of the most useful studies in epidemiology, and the revealed preference (observed behavior) data of people leaving their homes, etc., which provides objective measures of what would otherwise be subjective phenomena. In general, this is an exposure-disease combination where causation can be inferred from a smaller number of less formal observations than is possible for cases such as chemical exposure and cancer risk.

- The reported health effects, including insomnia, loss of concentration, anxiety, and general psychological distress are as real as physical ailments, and are part of accepted modern definitions of individual and public health. While such ailments are sometimes more difficult to study, they probably account for more of the total burden of morbidity in Western countries than do strictly physical diseases. It is true that there is no bright line between these diseases and less intense similar problems that would not usually be called a disease, this is a case for taking the less intense versions of the problems more seriously in making policy decisions, not to ignore the serious diseases.

- Existing evidence is not sufficient to make several important quantifications, including what portion of the population is susceptible to the health effects from particular exposures, how much total health impact wind turbines have, and the magnitude of exposure needed to cause substantial risk of important health effects. However, these are questions that could be answered if some resources were devoted to finding the answer. It is not necessary to proceed with siting so that more data can accumulate, since there is enough data now if it were gathered and analyzed.

- The reports that claim that there is no evidence of health effects are based on a very simplistic understanding of epidemiology and self-serving definitions of what does not count as evidence. Though those reports probably seem convincing prima facie, they do not represent proper scientific reasoning, and in some cases the conclusions of those reports do not even match their own analysis.

Personal background/credentials

My name is Carl V. Phillips. I am an expert in epidemiology and related health sciences, as well as scientific epistemology and methodology, and have been retained by attorney Edward Marion, representing Brown County Citizens for Responsible Wind Energy, to provide analysis and testimony in connection with Public Service Commission of Wisconsin docket no. 1-AC-231, Wind Siting Rules.

I earned a PhD in public policy (with an emphasis on economics-based decision making) from Harvard University, completing a dissertation on environmental policy and economics. I then completed the Robert Wood Johnson Foundation Scholars in Health Policy Research postdoctoral fellowship at the University of Michigan. Later I did a second fellowship in philosophy of science at the University of Minnesota. Before I returned to school for my PhD and began my career in public health science, I worked in consulting, primarily analyzing energy and environmental policy issues. Prior to that I earned a Master's in Public Policy from the Kennedy School of Government at Harvard, and *summa cum laude* undergraduate degrees in math and history from Ohio State University.

I spent most of my career as a professor of public health. I currently direct an independent academic-style research institute (a continuation of my university research lab) and a small consultancy. During my career as a professor, I taught at the schools of public health at University of Minnesota, University of Texas, and University of Alberta (Canada), the evidence based medicine program at University of Texas medical school, the University of Alberta medical school, and Harvard's Kennedy School of Government. My teaching focused on two subjects: how to make optimal public policy decisions based on scientific evidence and how to properly analyze epidemiologic data. This subject matter, as important as it is, is generally overlooked in health science and medical education, and students frequently reported that my teaching clarified their understanding of epidemiology, science more generally, and policy decision making for the first time in their educational careers.

My research during my academic career, and continuing in my private institute, has emphasized epidemiologic methods, environmental health, science- and ethics-based policy making, the nature and quality of peer review, and tobacco harm reduction (the main focus of the current institute). My work on epidemiologic methods focuses on recognizing and quantifying uncertainty, recognizing and correcting for biased analyses, and translating statistical results into decision-relevant information. My initial contributions in the area of quantifying uncertainty won several awards in the early 2000s and launched a new area of inquiry in the field.

Epidemiology is the study of actual health outcomes in people, and thus is the only science that can directly inform us about actual health risks from real-world exposures. Related biological and physical sciences often provide useful information about health risks, but they are ultimately trumped by epidemiology because real-world exposures and the human body and mind are so complex that we cannot effectively predict and measure health effects except by studying people and their exposures directly. My background in epidemiology methods, scientific epistemology,

and optimal policy decision-making is the complete background that is needed for being able to evaluate bodies of health science literature and assess their worldly implications. Most people who work in or around epidemiology learn only how to conduct particular types of studies or how to technically interpret individual study conclusions in the simplest possible way, which does not provide the ability to sort out complicated controversies. My study and research have focused on the epistemology of epidemiology, how to understand what the available evidence tells us beyond what the authors of individual studies assert.

My background in environmental economics and environmental health, with an emphasis on energy policy, provides important subject-matter literacy for the current case. I have reviewed the available scientific literature relating to the present case, including previous filings and summary reports and references included therein, as well as other information I felt necessary to find in order to form my opinions.

Summary of my scientific opinions

My main expert opinion in this case is that there is ample scientific evidence in support of the claim that wind turbines sited near residences cause substantial and important health problems for some people living in those residences. This does not mean that we have definitive evidence that there is a problem, but it would clearly be absurd to deny the possibility that there is a substantial problem. Further research could produce results that cause us to change our assessment, but the best assessment based on current evidence is that there is a problem, potentially of great magnitude for the exposed populations.

There is not sufficient evidence to robustly estimate what portion of the exposed population will suffer health problems of a given magnitude (though there is some evidence to allow rough estimates), or to assess much detail about the varying effects based on intensity of exposure. The exact list of diseases being caused by the exposure is difficult to determine, and there is little information to assess whether neurological damage or other physically measurable diseases exist. Nor is there sufficient evidence to assess exactly which of the several candidate causal pathways leads from the existence of the wind turbine to the health outcomes (which bodily systems; primary medium of transmission; nature of the noise and circumstances), and so it is difficult to assess the options to mitigate the effects.

In particular, my scientific analysis is based on the following points, which are expanded upon below:

1. Health effects from the turbine noise are biologically plausible based on what is known of the physics and from other exposures.

2. There is substantial evidence that suggests that some people exposed to wind turbines are suffering psychological distress and related harm from their exposure. These outcomes warrant

the label “health effects” or “disease” by most accepted definitions, though arguments about this are merely a matter of semantics and cannot change the degree of harm suffered.

3. The various attempts to dismiss the evidence that supports point 2 appears to be based on a combination of misunderstanding of epidemiologic science and semantic games. Multiple components of this point appear below.

4. Epidemiology studies could be designed and carried out to provide a much more useful assessment of the existence and quantification of the health impacts of wind turbine facilities.

Biological plausibility

When conducting epidemiologic assessments, it is useful to establish biological plausibility (i.e., that there is an apparent mechanism via which the exposure *could* cause the outcome in question). Such plausibility is not always necessary, and in a some cases major discoveries were made before any mechanism was recognized. But often when an association is observed without identifying any plausible mechanism before looking at the data, it turns out to be a statistical accident or a bias in the analysis. Identifying mechanisms via which an exposure could cause such human health effects does not prove that the particular mechanism is the pathway, nor, of course, does it mean that there are actual health effects.

In the case of the effects of wind turbines, the nature of the observed diseases means that plausible pathways are difficult to identify with certainty. Unlike a case where a particular physical exposure can be shown to lead to a disease, exactly what bodily process to look for is unclear, and there are multiple proposed candidates. But it is clear that the physical effects (noise traveling through the air, noise/vibrations traveling through the ground, and possible flickering light – hereafter just abbreviated as “noise” or just the “exposure”) do reach local residents, and noise does cause health problems under some circumstances.

More important (and presumably the impetus for studying the phenomenon initially) is that people can observe that the noise from the turbines seems to be bothering them, and can surmise that what they are noticing may be causing their diseases (the significance of this point is expanded upon below). It is well established that local residents can hear sound from the turbines under many circumstances and experience flickering light when the line from the sun to a house passes through the range of the blades. Apparently without exception, analyses of the physics of noise from turbines show that the noise reaches houses (though some who report such results argue that it does so only at magnitudes that cannot cause health effects).

Additionally, various studies show that lower level sound and vibrations affect the body via the ear (hearing and vestibular systems), skull, skin, viscera, and other body parts. The favored hypothesis among some who advocate for recognition of the health problems from turbines seems to be effects on the vestibular system. Other authors favor a stress-hormone mediated pathway. Identifying the existence of these possible effects on the body does not mean that the

exposure necessarily causes harm, but it does provide a plausible pathway for the types of harms that are typically caused by accumulated stressors.

Apart from studies of hearing damage, which does not seem to be at issue here, there are several lines of research that indicate that long-term exposure to noise causes health problems. While the literature is mixed, there are enough studies that suggest different health problems result from chronic noise exposure, at least under some circumstances, that it seems very difficult to rule out a particular problem out of hand. One group of researchers has focused on a problem they label vibroacoustic disease, in which substantial tissue effects are observed after long periods of noise-exposed employment (e.g., Alves-Pereira and Castelo Branco, 2004). While it does not appear that the conditions that produce that disease exist in the present case, the existence of such a disease is suggestive of the many not yet understood ways in which long-term exposure to noise can affect the body.

A common result of many studies is that exposure to noise that is not immediately harmful can, over a period of days or months, affect the body in ways other than simply hearing it. My non-systematic review of articles about noise and public health (those wishing to browse that literature can start with the journal *Noise & Health*) suggests that most of them deal with exposures that do not cause acute damage, and few offer a confident assessment of the causal pathways. With the exception of damage to hearing, noise causes problems via mechanisms that we do not fully understand. The significance of this failure to understand the pathways is important: It does not mean we should doubt there are causal relationships just because we cannot figure out exactly how they work. Rather, it means we should be open to the possibility that there are causal pathways beyond our present understanding. For example, there was evidence that immune system failure and oral cancer might be sexually transmitted based on the population statistics alone long before we recognized that the causal pathway from sexual activity to the outcome was a sexually transmitted virus (HIV and HPV, respectively).

<u>Evidence that individuals are suffering disease from wind turbines</u>

There are numerous reports of individuals residing near turbines suffering a collection of similar (across individuals) and plausibly related (as part of the same disease syndrome in an individual) symptoms, including psychological symptoms of general distress (irritability, etc.), failure to concentrate, and depression, as well as symptoms at the border of the psychological and physical, including insomnia, fatigue, headache, and tinnitus. The reported effects are typically described in terms of symptoms, the symptoms being both the sign of the disease and the source of suffering. When such patterns of symptoms occur briefly or are secondary to some more dire (or simply more concretely defined) condition, we do not generally think of them as disease in themselves. However, these can represent a major reduction in well-being and in health even if there are no other known effects. It is possible that some of the problems, particularly if experienced for prolonged periods, could lead to mortality-threatening disease via their effect on blood pressure or other pathways, but it is unlikely such downstream physical manifestations of the stress will be detectable (they will disappear into the random background rates and occur at an unknown time in the future), and so failure to detect them tells us little. But this inability to

measure, and perhaps even to establish, the worst-case manifestations does not diminish the importance of what we can establish and could measure.

There is a small collection of epidemiologic studies of people exposed to wind turbine noise in Europe by Pedersen and colleagues (2004, 2007, 2009, 2010). These studies suggest that some substantial portion of exposed individuals experience harms, some of which constitutes health problems by any modern definitions of health. The studies have various limitations, but they provide a quantification of a nontrivial number of cases. Phipps (2007) also presents a systematic study (which has been misidentified as a case series) that reveals similar results.

Other studies conducted less systematic analyses of exposed populations, providing case series or case studies. Perhaps the most prominent case studies in North America are those reported by Pierpont in her book (2009) and elsewhere. While it has been (correctly) noted that her work was not systematic in several senses, and thus cannot answer some questions of interest, this does not diminish its contribution as a report that these problems exist. Other case reports can be found in various collections (e.g., Harry 2007) and the self reporting of individuals who have wanted to report their own experiences. The latter are completely unsystematic and represent the most biased possible selection, but for this particular disease (for reasons expanded upon below) this does not mean that they provide no useful evidence about the existence of a problem.

Some recent commentators (Colby et al. 2009; Roberts and Roberts 2009) have attempted to dismiss this evidence because none of it is based on the epidemiologic study types that they understand. It is true that other study designs would have told us more, and still could. But dismissing the evidence we have makes little sense given that a huge portion of all knowledge, including formal scientific inference, is based on data that is not from studies designed according to certain preferred approaches. It should be obvious that "does not tell us everything we want to know" does not mean "has no information content". Those making this argument either do not understand scientific inference or are pretending they do not. Claiming that there is no evidence even though there are reports of individuals suffering is akin to claiming that there is no evidence that people get injured as a result of text-messaging while engaged in other activities because, even though the pathway is obvious and there are numerous accidents occurring from some activities, there is often not a "real study" that allows us to make various quantitative estimates.

The most fundamental flaw in that reasoning is that some of the case studies of those exposed to turbine noise actually constitute a quite useful study design, albeit accidentally. The *case-crossover study* (Maclure 1991) is one of the most effective methods for assessing the transitory effect of a transitory exposure (which means in this context: occurring fairly soon after being exposed to turbine noise and disappearing fairly soon once the exposure is removed). This study design is one familiar to all of us in our everyday lives: Impose the exposure at a time when the outcome of interest is absent and see if the outcome occurs; withdraw the exposure and see if the outcome disappears. For example, if you think that a particular food gives you stomach pain, note whether that pain occurs after eating the food but not at other times; avoid eating the food for a while and see if the pains are absent. It is obvious how to translate the resulting

observations into a causal conclusion about the food and the pain. For an even simpler illustration, our usual study design to figure out if a particular switch turns on a particular light is to flip the switch (often three times for some reason) – nothing more complicated is needed. When it is possible to collect case-crossover data, it usually provides among the best possible epidemiologic information. Its advantages include more data (each exposure change serves as an observation, whereas most other study designs produce only one observation per person) and individuals serving as their own comparison population.

For the case of turbines, many of the reports (personal testimonials and collected case series) recount the onset of distress being shortly after with the activation of the turbines – that is, when the person crossed over from being unexposed to being exposed. Moreover, some residents have reported reductions in their health problems under certain conditions (based on wind direction or speed), further supporting the conclusion. Finally, some of those who sought relief from their symptoms through physical (soundproofing) or geographic (moving their home) methods crossed back to unexposed and presumably achieved relief from the disease. While the cause-and-effect pattern might not be so obvious as the light switch experiment (because the appearance and disappearance of symptoms is not quite so immediate), this is fairly compelling evidence.

Several of the case studies reporting residents' exposure to turbines hint at exactly such case-crossover data. Unfortunately, it does not appear that anyone working in this field previously understood the full potential of the case-crossover study, and so did not collect all the possible data about timing and crossing back in the optimal form. But many of those reporting case studies seem to have intuitively understood the importance of observed crossover contrasts (as most people do, with or without formal scientific training). Examples can be found in Pierpont's reports, for example, and it might be possible to mine previously reported case studies for more formal presentation of case-crossover evidence. Failing that, it is relatively quick and easy to collect such data from those already experiencing problems from existing facilities, especially if previous researchers have maintained contact information for their subjects.

The second crucial epistemic consideration is that this exposure-disease combination is quite different from what is typically claimed for noxious facilities. In cases of environmental pollution there is often a fear of slow-developing diseases (especially cancer, for which we cannot even define a time of incidence – i.e., when the disease per se actually started – only of diagnosis) that occur seemingly at random because they have many causes that are impossible to sort out. In such cases, when local residents claim "I got cancer because of the effluent from this factory" the standard response is that it was inevitable that some people near the factory would get some cancer someday, and so it is impossible to make that causal conclusion. To make any such conclusion it is generally necessary to systematically collect enough data on enough exposed cases, as well as on non-cases and an unexposed comparison group, so that statistical comparisons can be made. (The caveat "generally" is meant to recognize the fact that if twenty cases of the same rare cancer were reported among a few hundred exposed individuals we would not actually need to know much more to draw the causal conclusion.) Contrast this epistemic situation with the case of a traumatic injury from a car crash: If following a crash a passenger in

the car has a laceration on his head that he did not have a few minutes earlier, we would not hesitate to say, based on that information alone, that the crash caused the injury. Why? Because head lacerations do not slowly develop from unknown causes, appearing perhaps years later (like cancer); instead they are almost always diagnosed shortly after a causal event occurs. Moreover, we can almost see the causal pathway in the form of the crash causing rapid deceleration which caused an impact between head and something in the car, and it is such impacts that cause trauma.

The case of the various forms of distress caused by wind turbines lies somewhere between the cancer and crash examples, but is rather closer to the latter. While there is contrast with the trauma case in that we cannot see the causal pathway, the particular form of distress that has been observed is not something that often suddenly occurs without some observable proximate cause. This contrasts with cancer caused by chemical exposure, where the chemical insult that triggered it is often invisible and possibly long in the past. The reasonable expectation of a proximate explanation and ability to observe the turbine noise as the ostensible cause make this case more like the car crash than the cancer. Indeed, some commentators who have sought to dispute the claim of health effects from turbine noise have proposed the theory that those suffering health effects after local turbines began operation would have just lost a battle against the siting of the turbines and would be suffering from exhaustion or a sense of defeat from that, and those factors would bring on the symptoms. Others have suggested, somewhat more plausibly, that the cause of the effects is that publicity and local fear about health risks ironically creates the distress. Whatever the merits of the arguments (they could be evaluated with research rather than just asserted), they implicitly acknowledge that most observers agree that onset of the observed health problems generally has an explanation in terms of a recent event.

Moreover, though it is not quite so easy to observe the cause of distress and the resulting psychological and physiological manifestations as it is with the crash and head trauma, a subject's own observations about his own single case are still scientifically informative. This contrasts with most types of cancer, wherein neither the victim nor any clinician or scientist can offer a legitimate conclusion about causation, other than in the form of far-from-certain probabilities derived from statistical comparisons. Someone who claims "this exposure caused my (or this patient's) cancer" is never justified (except for the few nearly single-cause cancers). But if someone claims "this noise is driving me crazy and keeps me from sleeping" we have good reason to believe him. For a more subtle exposure, like low decibel low-frequency noise, the conclusion is less certain than it would be for a loud party next door, but the individual's assessment still has substantial value. This is true even apart from the crossover data that an individual will naturally accumulate (e.g., by spending time away, by changing rooms and other mitigation behavior, etc.), and with that crossover data and common intuition about how to use it the conclusion is even more definitive.

Thus, unlike the case of trying to detect an elevation above some baseline level of a disease that usually has very distant and uncertain causes, which is usually impossible absent formal studies that are designed to do just that, the natural observations in this case are quite compelling.

[Sidebar: This does not appear to have been brought up in the documents associated with this matter, but it might be, so it is worth immunizing readers against a particular common error: mistaking prevalent cases for incident cases. It is imaginable that someone might respond to the points presented here by saying "at any given time, many people are suffering from the collection of symptoms discussed in this context, and most of them are not exposed to wind turbines". That author might go on to argue that therefore we should expect to find many exposed cases purely by coincidence. But such a statement describes *prevalent* cases – i.e., cases that exist at a given point in time – rather than what we observe with turbine exposure, *incident* cases – i.e., cases that begin during the observation period. The distinction is important because many prevalent cases are long-existing diseases which should not be confused with a case that appears shortly after the exposure begins; coincidental cases that become incident the week or month turbine operation begins will be exceeding rare. This does not mean that they might not have some other cause, like the fear or political battle hypothesis that were just mentioned, but mere coincidence is unlikely. Measuring the coincidental occurrence of exposure and disease together, and then assessing whether there are extra cases that are caused by the exposure, is the purpose of most of the complicated statistics used in epidemiology. But the number of coincidental incident cases is very low for the car crash or turbine examples, making the complicated statistics less necessary to merely establish that something occurs (though they are still needed to quantify the risk). But someone who does not understand the difference between prevalent and incident cases might not recognize this.]

Observations about behavior and prices can further support the claim of a causal relationship between turbines and health problems, as well as demonstrate the great magnitude of some of the problems. An apparently nontrivial portion of the residents whose cases where reported in detail, people who had concluded that the turbines were damaging their health, have moved their residence or retrofitted the structure to try to block the noise. These are expensive actions that would not be taken by people who were suffering only minor problems or who had not made every effort to make sure the cause of their disease was indeed the turbines.

Some of the residents attempted to sell their properties and failed to find any buyers at a price they would accept, suggesting that potential buyers anticipated suffering the same problems if they moved near the facility. It might be surprising to see such observations being used as epidemiologic data, but their value should be immediately clear to anyone trying to carry out inference, and who understands more than how to follow a simplistic recipe for how epidemiologic analyses are done. It is theoretically possible that everyone involved (residents, all potential buyers) is so misled about the causes of their health problems that they would upend their lives or waste thousands of dollars in error, and others would fail to take advantage of their error by buying their heavily discounted houses, but economists recognize when there is data like this (called "revealed preference"), it is usually the most compelling evidence available. Of course, there is likely a bias in that those suffering the worst problems are most likely to report their experiences. So once again, case study evidence is not adequate for quantification. But a

systematic study of actions by residents and of the real estate market could offer very useful information.

Summary of what can be concluded from the affirmative evidence

Looking at just the affirmative evidence, then, there is much to support the claim that some people suffer substantial health problems as the result of the externalities from the turbines. At the very least, before considering counter-arguments, any unbiased observer would have to conclude that, at a minimum, substantial health effects are quite plausible, seem to have occurred, and certainly cannot be ruled out without further directed study. To estimate the portion of exposed individuals who experience these problems would also require a more systematic study, as would quantifying the total health impact throughout a population.

The following several sections respond to some of the counter-arguments that have been offered to the conclusion that the existing evidence strongly supports the claim that there is an effect. Many of these are simply attempts to distract readers and do not represent any form of legitimate scientific argument, and so are mentioned briefly to call attention to their lack of legitimacy. Others require a bit more analysis.

Detection or measurement of physiological effects etc. are not necessary for there to be evidence of health effects

Claims have been made that people cannot usually detect the sounds or vibrations from wind turbines at certain distances (e.g., Colby et al. 2009 address this point multiple times and seem to imply it is of great importance), though there are ample counterclaims about people being able to detect the noise. The claims, whether or not they are accurate, have no bearing on whether these exposures affect people's health and well being. Many things we cannot detect can harm us. Though it is not clear that anyone is saying this outright, there is rhetoric that seems to say "the causal pathway does not pass through the most obvious effect of noise – hearing or otherwise detecting it – and this constitutes evidence that there is no causal pathway." But simply eliminating one of the many possible pathways tells us nothing about whether there are health effects, especially when the actual health effects have been observed.

(Aside: For those not used to reading serious analyses of scientific epistemology, it is important to not mistake "even if this were true…" for "this is true" when reading this document. The argument in the previous paragraph is that *even if* we assume there is no detection, it would not be evidence that there is no problem and so that argument is simply unscientific and thus, for immediate purposes, it does not matter whether there is detection. But later, actually making the "even if" assumption could impede further the empirical work or remedies, so it should be recognized only as an inconsequential assumption for a particular purpose, not all purposes. Additionally, *even if*-based arguments should not be interpreted as meaning that the author believes that the claim is true.)

The lack of observable physiological effects associated with (and, in particular, apparently causing) the observed psychological distress is also not a basis for dismissing the evidence we have. The observation of the lack of a known physiologic mechanism may be intended to argue that "the physical phenomenon of noise would have to initially affect people physically [which is true], and we do not fully understand what this effect is [also true], so the health problems are probably not real [a completely unsupported conclusion]". A better conclusion from our ignorance is "it could be useful to do more research to figure out what causes the problems". After all, we do not even fully understand exactly what physical impacts from smoking cause all the heart disease or even lung cancer.

The dismissal of the evidence is sometimes so bald that it seems like parody. Colby et al. (section 4.1.2) go so far as to write "There is no evidence that sound at the levels from wind turbines as heard in residences will cause direct physiological effects. A small number of sensitive people, however, may be stressed by the sound and suffer sleep disturbances." Even if the latter characterization did not comically understate the evidence, these authors, within the space of a two-sentence paragraph, claim there are no physiological effects but note that there are observed cases of turbines causing a physical problem. (One can speculate as to how the authors rationalize this. Perhaps because sleep disturbances may be a manifestation of what are primarily psychological effects they are not counting them, but since they certainly are a physiological manifestation, this makes little sense. Perhaps the authors are trying to gerrymander "physiological" to exclude physical effects that can be measured outside a laboratory; it is not even sufficient to torture the definition to only include effects that can be measured by biomarkers, since sleep loss can be measured by biomarkers if necessary. Or perhaps the crux of the game is found in the word "direct", which is really nonsense since additional intermediate steps can always be inserted into a causal pathway, so the word is inherently meaningless in this context. Whatever the authors thought was a sufficient rationalization, it is clear that they are making great effort to rationalize denying the obvious conclusion, that there is evidence of physiologic effects.)

<u>Psychological diseases are real and very important from a public health perspective</u>
Some commentators have tried to dismiss the reality of the reported diseases because they are primarily psychological. Psychological diseases and those with manifestations on both sides of the psych-physical border – a category that includes stress, depression, and many other ailments – arguably account for the loss of more quality-adjusted life years than purely physical diseases, at least in the West and possibly even worldwide. Most all accepted definitions of individual or public health include psychological health as part of the consideration, and usually refer to an overall state of well being rather than just an absence of particular diagnosed pathology.

It is sometimes tempting for people who do not suffer from addictions, depression, or other psychological or physical-and-psychological diseases to have the attitude that sufferers do not have a real disease and should just snap out of it. There is a substantial literature that documents this dismissive attitude (which in the present case might translate into "sure, I can hear the noise if I think about it, but it is no big deal, so it should not really bother you either"), and argues that

this makes no more sense than telling someone with appendicitis that he is just being negative and should snap out of it. For example, it is difficult for most of us to imagine why someone would have the urge to drink himself into disfunction every day and not even apparently enjoy it, and so tempting to think "all he has to do is realize that this is a bad thing to do and he will quit". But it should be obvious that this is not sufficient and his affliction, however difficult it might be for most of us to empathize with, is quite real.

Some of the attempts to dismiss the importance of the observed health problems are semantic games and belittlement, cheap tactics that are typically used to obscure the lack of legitimate scientific arguments. The term "annoyance" has been adopted as jargon by researchers to refer to certain psychological problems resulting from noise, and has the implication (presumably intended in some cases and not in others) that the causal pathway from noise to disease involves the type of psychological experience that is typically referred to as "annoyance" in natural language. However, it is clear that there is nothing in the use of the jargon that implies that the harm is minor, and indeed the term is used in contexts in which it includes life-destroying harms. But some of those who seek to suggest this suffering is minor appear to be trying to confuse readers based on the natural-language meaning of the word, since in natural language, "annoyance" typically implies a minor harm, a "mere annoyance" that is well below the level that would be considered a disease. But of course adopted jargon does not have its original natural language meaning. Just as "insult" can mean cause of cancer, "annoyance" in this case includes serious physical and psychological symptoms.

Language games like these tend to suggest an attempt to avoid direct discussion of the evidence that there really is a problem: If a writer wishes to claim that the various symptoms are not actually a health problem, he should say so (and explain why most definitions of health say otherwise); if he wishes to say that the health problem is so minor that it should not be called a disease, he should say that (and explain why someone would leave their home over a minor problem). Referring to the effects as "just" annoyance does not constitute an argument that the effects are not disease nor that they are minor.

Similarly, some commentators (see Roberts and Roberts, p.39) have attempted to belittle the disease experience with the phrase "dis-ease", which they implicitly define as including pain, anxiety or physical manifestations thereof. They concede that their construct can often be "undistinguishable from the state of disease as related to morbidity" and then try to argue that it is somehow fundamentally different (rather than reaching the more obvious conclusion that two things that are indistinguishable perhaps should not be distinguished when making policy decisions). They argue that "with physical illnesses, objective measureable [sic] indicators can be obtained through instrumentation testing that is typically absent of human error or influence." This is absurd, both because not all physical illnesses have objective measures (e.g., headache and other pain-defined diseases) and because nothing is ever absent of human error or influence. They then point out that, "Subjective responses to stimuli are much harder to prove or disprove which is why it is very important to supplement a subjective response with an objective assessment", which is reasonable advice, though scientists know that you make the best of

whatever form your data takes rather than declaring the data to be less than one might want and pretending complete ignorance. But they then go on to ignore the non-subjective data that exists, as pointed out below. The rhetoric those authors use is hard to interpret as anything other than an attempt to distract from the evidence.

Another scientifically-invalid bit of rhetoric is to observe that the observed collection of health problems is not a single officially-designated disease, and imply that this makes it somehow less real. Setting aside whether it would be useful for simply parsimony to have something officially defined as "wind turbine syndrome", it is clearly not necessary to have such for there to be a real collection of health problems caused by turbine exposure. The title would not make the phenomenon any more real. Indeed, syndrome-based diseases with primarily psychological symptoms are generally defined based on a systematized (and often only slightly systematized) version of "you will know it when you see it", and definitions are created, altered and dropped over time based on both science and politics. While this does not mean that such diseases are any less important to people's health, it makes clear that the lack of some official designation is fairly meaningless. After all, there is no official disease designation for "injuries resulting from a driver of a subcompact car being hit from the left side by an SUV", but there is certainly a collection of injuries that typically occur from that, and anyone could give it a label.

The most legitimate point that can be mined from claims about the "realness" of the disease is that there are some challenges in dealing with diseases that are subjectively measured (i.e., can only be ascertained and quantified based on the victim's own assessment), especially in a politicized situation. Observing that there is this challenge does not make the diseases non-real. Many important diseases are diagnosed and largely defined based entirely on subjective experience, everything from suicidal depression to a minor headache. To dismiss subjective experiences would be to dismiss the vast majority of what people genuinely care about in the world, as well as many fields of science and medicine. But there are challenges that must be considered.

One challenge in assessing the importance of subjective sources of distress, pain, etc. is that context matters: A minor trauma might cause a pain we just shake off while playing sports but the pain from such a trauma might be completely unpalatable if it occurs while sitting quietly at dinner, and if that pain were inflicted by someone walking by and hitting us on the head we might consider it even more painful and a criminal act. Similarly someone working in a job with a sense of élan (such as training to be an astronaut, as in one of the studies about the effects of noise, but probably for any job someone is happy with or proud of) will likely be more like the sportsman than the person sitting at dinner, and might intentionally downplay the pain or distress. Thus, occupational studies of the effect of noise would be expected to show effects different from – likely less than – those from residential exposure. The challenge is not just that the experiences are different based on circumstance, but that neither is more legitimate than the other. If the subjective experience of the resident at home is different from the worker on the job, it is no less real as a result – a scientific or ethical argument that we should dismiss it for this reason would require ignoring most all pain and similar experiences. Observing that sometimes

people are able to ignore pain or another subjective effect might suggest intervention methods for reducing the impact, but the observation does not in itself reduce the impact.

Sometimes the commentators who seek to dismiss the importance of the health problems appear to misuse "subjective" to mean "psychologically mediated", and thus include all the symptoms typically discussed in the context of turbines. Some of the disease outcomes that have been cited as being caused by turbine exposure are not actually subjective. Loss of sleep, for example, is objectively measured even when measured by an individual (even though he is also the subject, the method of observation – counting up hours that he was asleep – makes him an objective observer of himself). It is actually even possible to objectively measure pain and related sensations – we do so whenever we observe sentient beings other than ourselves endeavor to move away from some stimulus. No one ever knows exactly what another individual is experiencing (that is the essence of subjective symptoms), but we can see when they are reacting in a way that provides convincing objective evidence of distress. Someone moving their home certainly fits this description, a subjective experience with an objectively observable manifestation.

There is no clear way to draw the line between health problems per se and other negative effects that should not be classified as disease. Most everyone would agree that a momentary noise that awakens someone for a few minutes imposes a real cost but the experience does not constitute a disease, while a constant noise that consistently prevents sleep for weeks or months, leading to depression and the many effects of insomnia, has created disease. The point in between that marks the border between disease and other types of costs is arbitrary and not defined. The simplest response to this is that because the difference is arbitrary, it really should not matter for any practical purpose. Inflicting harm on someone is equally unacceptable whether that harm takes the form of "disease" or not. Forcing someone to spend 100 hours dealing with miserable legal proceedings is not appreciably different from causing them to suffer 100 hours of severe flu symptoms.

Politically, it is sometimes the case that harms are considered more important if they can be designated as diseases. For example, instead of recognizing that the aesthetic and minor immediate health objections to involuntary exposure to second-hand smoke can justify bans in indoor public spaces, bans have only been implemented based on claims about life-threatening disease risk. In the present case, many of the above points – whether there is a "wind turbine syndrome", and even whether the effects are disease, "dis-ease" or "annoyance" – are based on the assumption that a harm does not matter so long as it does not have the label "disease". But not only is there no bright line between suffering certain types of diseases and otherwise being forced to suffer a less pleasant life due to ailments, but the same sciences (epidemiology, econometrics) that measure the former inevitably also measure the latter. Certainly from an economics or policy ethics point of view, there is no meaningful difference: A major cost inflicted upon someone's psychological well-being does matter, whatever it is called. Anyone who is attempting to argue that the harms do not represent an official disease and that this affects how we should treat them, should be asked to declare explicitly (as they are admitting implicitly)

that there really are effects that people care about, but they simply should not be called diseases. Only having done this is it proper for them to argue that these effects should be ignored in policy decisions because they are not disease, which is different from arguing that they do not exist or are small.

Heterogeneous effects are to be expected

Some observers might be confused by the fact that some people apparently experience debilitating symptoms from their exposure to turbines, while others may have greater exposure (as measured physically) but no significant symptoms. But this is not at all unusual, and similar patterns can be observed for most any exposure. For example, many heavy smokers never get cancer or suffer any other major disease that is often caused by smoking. There are hypotheses and some data about who is more likely to suffer from exposure to turbine noise, but this could be better informed by further study.

Some observers have had brief exposures to the noise and experienced no adverse effects, and perhaps concluded that the exposure would never bother them. But this does not constitute evidence that no one ever suffers from the effects: The individual in question might be immune while others are not. Or he might be wrong about what would happen if he were exposed longer, since health problems caused by noise exposure tend to be cumulative, as is typical for other exposures that produce stress reactions (social harassment, pain, sleep deprivation, physical restraint) which may seem trivial for an hour but torture after a week. Even exposures for eight hours per day (like workplace exposures) may have quite different effects than exposures that last all day and overnight. Some exposures that people intentionally seek for an hour or a workday-length period (hot weather, loud music, exhausting exercise) cause stress reactions and health problems with unrelenting long-term exposure. The use of low frequency sound as a method of therapy, which at least one group of commentators tried to portray as evidence that such sound is always harmless (Colby et al. 1999, p. 3-17), is actually further evidence that these exposures are difficult to analyze other than via epidemiology because the real-world "dosage" of the exposure matters a great deal (those exposed to the turbines are presumably not sharing in the therapeutic benefits). Thus, observations about the limited effect on some people or of short-term exposure cannot be seen as denying serious effects on some people who experience long-term exposures.

Colby et al. emphasize the empirical observation that the effects of turbines on people depend on personal characteristics (p. 5-2). They do not explain why, but they seem to either find this surprising or want to imply that it is an argument that the effects are less "real". Either of these explanations would suggest that they understand little about epidemiology (a conclusion that is supported in depth below), since anyone familiar with the science knows that the effect of every exposure varies with personal characteristics. Entire sections of the science are dedicated to figuring out how to optimally deal with this fact. Those authors seem to be making the mistake, common among people who do not understand complicated sciences like epidemiology, of thinking that if an outcome has one cause (personal characteristic) then another factor (noise) is not really the cause. In fact, the proper way to think of it (though it takes some getting used to

for non-scientists) is that *both* the noise and the personal characteristics caused the disease (as did a multitude of other factors); if either one of them was absent then the disease would not have occurred. In ethical or policy discourse (as opposed to scientific analysis) we often reduce our list to causes that someone actively brought about (i.e., the causes someone is culpable for, not the ones that simply are). So, for example, a murder is not excused, and is not considered to not be the cause of the death, if his victim would have survived had he been stronger or closer to a hospital – those non-act-based personal and geographic characteristics *also* caused the death, but the murderer is still a cause and thus is guilty.

Health problems are related to negative opinions about the turbines

Some commentators have made the observation that there seems to be a correlation between health problems and a negative opinion about the facilities (see, in particular, Pedersen and Waye, 2004), and have insinuated that the health problems are therefore less real, or perhaps even concocted due to other motives for disliking the facilities. Exactly why even a local resident who disliked the facility would concoct or exaggerate health effects is unclear, given that it is exceedingly unlikely that they could cause any existing facility to be removed, and this has not been explicitly alleged, so I will not address it further.

We should obviously expect to see the observed correlation when data is collected after the turbines are operating (which includes cases where people proactively report their experiences with health problems): Anyone suffering new health problems that they perceive to be caused by the turbines is going to have a negative opinion. The health problems cause the dislike of the facilities, which manifests in hating the sight of them, etc., not the other way around.

Even when disposition data is collected before the turbines start operating, there is still a good chance of causation running from health concerns to disposition. People who recognize, from experience or other self awareness, that they are more likely to suffer health effects from noise pollution are among those who will most strongly object to the siting and have negative feelings about it. Indeed, it seems safe to predict that a larger than average portion of the population with those feelings will be near new facility sites, since they local residents have chosen to live in quiet rural areas. It is certainly the case that local residents will be more sensitive, on average, than people who self-select into noisy occupations (i.e., the people who are the subject of most studies of the effects of noise).

This observation is related to the magnification of the health effects caused by the physical insult that results from fear and possibly frustration. Colby et al. discuss this at length, labeling it a "placebo" effect (and adding the silly neologism "nocebo effect"). Such labeling does not make the health effects any less real or devastating: A cure of a disease by a placebo is still a cure, though we do not fully understand why. In this case, the magnification of the harm due to fear and frustration is actually quite predictable. It does not make that harm less real or important from the individual or public health perspective – it just means that they could be addressed via different interventions. Indeed, if it is the belief of the industry or government that the substantial health effects that have been observed are due to "stress, fear, and

hypervigilance" (Colby et al., p. 4-4) rather than the physical effects of the turbines then they should be promoting interventions to eliminate these phantom ailments via education and counseling. (Note: just broadcasting to people "you are just imagining this and there is no real risk" is well known to have no effect, even when there is little scientific doubt it is true.) If the problems cannot be eliminated this way it would be evidence that either (a) the problems did not really have this "all in their head" characteristic or (b) there is no practical difference between these "placebo" effects and the other effects. The industry's failure to report on successfully intervening in this way suggest that they do not believe their consultants' rhetoric.

Forms of the evidence

Commentators who seek to deny particular health claims frequently resort to insisting there is "no evidence" because the evidence does not adhere to some criterion that they concoct. Liability defendants have demonstrated that it is almost always possible to argue that no study looked at exactly the circumstances of the plaintiff's experience, and therefore no evidence bears on the case. The present case, though not a matter of liability, is quite similar. If we disallow extrapolation of evidence from one situation (population, exact exposure) to another, then we simply have no scientific information about anything. As an exaggerated illustration, all studies took place in 2010 or before, so if we do not allow extrapolation it could be argued that we have no information about what will effect people's health in 2011 and thereafter. Legitimate scientific inference is a matter of figuring out what evidence shows about situations we have not observed, not looking for an excuse to claim that we know nothing relevant.

Sometimes the attack on the evidence takes the form of favoring some study designs over others. It might be that we wish we had a particular type of data, since some studies could tell us more than others. But failure to have the perfect data obviously does not mean we have no data. We simply need to be careful about only drawing the conclusions we can from the data (some people seem to be suffering) and avoid conclusions that are not possible to draw (x% of those exposed suffer some disease).

Another tactic for dismissing evidence is to argue that scientific analyses that are not in a peer reviewed journal are uninformative. It should be obvious that this might be a ground rule for a term paper or a debating club, but is not a good rule for truth-seeking. Very useful information can come in forms that are unlikely to make it into journals regardless of their information content (e.g., a compelling book of case studies will not fit into a journal article, but a barely relevant experiment on mice will likely get published, especially if it can be analyzed in a way that produces the result the researchers prefer). While there is some legitimate concern that enormous amounts of pseudo-scientific claims are written and we need a method for avoiding them, plenty of junk science appears in the peer reviewed health literature. Thus, while it might be that science in health journals is on average better than science that is not in journals, there is no bright line. Peer review does not promise accuracy, since reviewers can almost never assess the actual analysis (they do not have access to the data or the calculation methods, or even know the details of the methods (Heavner et al., 2009). And there are many kinds of useful peer review; the Pierpont (2009) book, the source of much argument in this area, appears to have been

peer reviewed more completely than most papers that appear in journals, and thus the arguments that it should not count because it was not peer reviewed represent either an ignorance of what the peer review process really is or pure rhetorical maneuvering.

Analyses of specific scientific reports

Most of what I have to say about the scientific and pseudo-scientific claims made in scientific reports is addressed above, addressing the substance of the claims. However, there are a few comments that are worth making outside of the context of specific scientific arguments. As I understand from what has been represented to me, and based on my reading of what appears in the literature, the main analyses prepared on behalf of the industry in this matter are the reports by Colby et al. and Roberts. Because these might tend to influence policy decisions, even apart from their specific arguments that are addressed above, I believe it is important to further illustrate the failings of these reports as legitimate health science analysis. While many analysts insist on only writing competing monologues that address opposing arguments only obliquely, if scientists do not directly and explicitly confront scientific errors, non-scientific readers generally have a difficult time determining which of the competing claims is accurate. Thus, it it part of the duty of those offering scientific advice for policy making to explicitly explain why claims that they disagree with lack validity.

It is notable that the Colby et al. panel did not include any population health researchers, even though the question they claim the report addresses is one of population health. Their expertise seems to be limited to the relevant physical sciences and clinical medicine. This explains the dominance of physics-based analyses in the report, discussions that are interesting and informative in some ways, but have very little bearing on question of actual health effects. From the perspective of a population health analysis, those entire sections can be summarized by the following: Turbines make noise; noise often affects people's bodies and health, though in the present case – if it can be established that there are health effects – we have no idea exactly what physical pathway leads from the turbines to the effect. Beyond that, these sections appear to be little more than general background about the technology. More cynically, they might be seen as impressive-sounding filler that might lead the causal reader to think there is a lot more substance to the report than there really is.

Some clinicians are trained and qualified population health scientists, but there is no evidence that those who participated in the Colby panel have such expertise. Unfortunately, it is often the case that physicians who may be quite skilled in their clinical abilities do not recognize that population science requires an entirely different set of skills. (Moreover, in the present case there is something quite curious to a population scientist who has years of experience observing physicians making population science conclusions without understanding the science: Usually physicians can be relied upon to say that if they have seen a case of a disease then the disease exists, and then they often err by over-concluding (about prevalence and especially cause) based on nonsystematic observations. In this case, however, they seem quite anxious to claim that we have seen the cases but they do not really exist, a very un-physician-like behavior.)

Section 4.5 of the Colby report exemplifies the lack of expertise in population health science. It is difficult to make this clear without seeming petty, but this section reads like it was written by someone who took a single class on how to understand epidemiology, and half understood the material. Like most readers of epidemiology, they present statistical significance as the measure of the accuracy of a study. Setting aside the complexities and common errors associated with this concept, they make a huge error: When they invoke it they are talking about a case series where there is no statistic that even could be statistically significant (or fail to be); the concept does not even apply. They go on to characterize "statistically significant" (in this context where it does not actually mean anything at all) as the opposite of the vague concept "simple coincidence" when it actually refers to a fairly technical test of how likely a particular observation is due to sampling error. But "coincidence" mostly invokes the situation where cases that occurred among the exposed population but not because of the exposure (as I use it above), a concept that is quite different from statistical significance or the lack thereof.

(For those who are interested, what the authors seemed to be trying to understand and explain is the contrast between associations (which cannot actually be calculated from the particular data) that have a low degree of statistical robustness – because they quite plausibly could have resulted from random sampling error – versus more robust results that we would not expect to see as a result of that random error alone (a particular technical definition of which is "statistical significance").)

The question of whether "there is a causal relationship between the exposure and the disease" (p. 4-12) is not a matter of whether there are control subjects, as they characterize. Also, they use the word "uncontrolled" to refer to lack of comparison populations, though this is a strange term to use for this (when used in epidemiology, it almost always means the analysis does not include covariates to try to control for confounding or is used to differentiate a purely observational study from a trial/experiment when the investigator controls the exposure). They describe case-control studies in the (admittedly common) naive way – that the comparison is between people with and without the disease, while that is merely the way the data are collected and like most other epidemiologic study designs the comparison is between those with and without the exposure. But they also offer the strange characterization that such studies always match subjects on other variables that might affect the probability of having a disease, which is one option for doing such a study but not the only one. They claim to describe a cohort study (which follows a population to look for new events) but actually use language that betters describes a cross-sectional study (a fundamentally different design which takes a snapshot in time looking for existing ongoing diseases). They imply that the only types of studies that exist are case-control and (the mis-described) cohort, ignoring cross-sectional, the usually favored randomized trial, the critically important case-crossover, and others.

These errors paint a picture of authors who are dramatically overstepping their expertise and hoping that no reader will ever have the expertise to notice and a forum like the present report in which to expose it. But even more important than these failures to understand epidemiologic methods, they fail to understand how to draw scientific conclusions in epidemiology.

In Colby et al.'s conclusions (section 5), even after citing many pieces of evidence that suggest turbines are having health effects, they repeat their "no evidence" claim. They apparently are basing this on the observation that "there is nothing unique about the sounds and vibrations emitted by wind turbines", combined with the claim that "the body of accumulated knowledge about sound and health is substantial". Their logic (they do not explain, so I must infer) seems to be "this is just like other noises" and "we already know everything there is to know about those other noises and they do not cause health problems". The first of these is utter nonsense. While their physics studies many not be able to identify what the relevant differences are, anyone who understands epidemiology knows that similar exposures sometimes have quite different effects. No other exposure is going to be exactly like the noise from wind turbines. Indeed, the "substantial" body of literature that they cite is not really all that impressive, and covers in depth only a few of the many forms of exposure to noise that people experience. In effect they are saying "we would not have predicted, based upon the limited analysis we can do using analogy and extrapolation, that health effects would have been observed, so we are going to insist that they really were not actually observed".

Perhaps this can be attributed to just sloppy presentation of summary points. But deeper flaws in their scientific reasoning can be found. They claim that some quota of studies proves a causal relationship while fewer tell us nothing. Their example is that "multiple case-control studies were necessary before the link between smoking and lung cancer could be proved" (p. 4-12), when in reality the first English language studies (which appeared almost simultaneously) are generally regarded as being quite sufficient for reaching the conclusion. (Moreover, the example is a very poor choice since – as an expert in epidemiology would know – those were the studies at the center of establishing the validity of the case-control study in the first place, so any uncertainty was more about the study design, not its results.) Sometimes a single study is quite convincing, while other times a collection of studies leaves a lot of room for doubt. This is contrary to their assertion that "only after multiple independent-controlled studies show consistent results is the association likely to be broadly accepted".

That last line is wrong at several levels. The first demonstrates their lack of understanding about what we are actually trying to infer (as well as the nature of epidemiology), since the question is not whether the association is accepted, but *causation.* The association is apparent in each individual study, or not. The phrase "independent-controlled" might appear to be jargon from the field, but it actually has no obvious meaning, while "broadly accepted" is a measure of public opinion, not scientific inference. Importantly, no studies ever prove causation – that is not how science ever works. So when Colby et al. denigrate case series data as not being able to "prove that an exposure is really harmful" (p. 4-12) they are saying nothing of substance. What they say that is of substance is that case series "can do no more than suggest hypotheses for further research" which is nonsense. To offer just one clarifying example, continuing on from a previous example, early case studies of tongue cancer cases in young people were sufficient to show that it was not being caused by tobacco use or drinking (on which most geriatric oral cancer was blamed), and later case series of oral cancer that tested cancer tissue for HPV

provided very convincing evidence that that virus was often the cause. Whether a particular study provides useful information about a question is not a simplistic function of the study type. In this case, since the most important question is "does it appear that turbines may be causing diseases", the case series is entirely adequate.

Indeed, this failure to understand what they are analyzing is worth emphasizing. Colby et al. write a (mangled) discourse on what is supposedly required to establish, "prove", or make "broadly accepted" a causal conclusion, when the conclusions of their report are basically that there is no evidence that there is any problem and no reason to do further study. In other words, they lay out (what they claim is) the burden to prove a hypothesis of a particular exposure-disease relationship is true, but then try to use the (claimed) failure as the basis for saying that the hypothesis is false. It is possible that this is calculated misdirection, though my reading of their many failures of scientific reasoning suggest that this elementary error may well be inadvertent. It is really difficult to believe that people who wrote that section have any understanding of epidemiology. And since epidemiology is the core science for understanding human health effects, it seem rather odd that this report is characterized as "an expert panel review" of "health effects".

Similarly, Roberts and Roberts (referred to for convenience as just "Roberts"), though specifically retained as epidemiology experts, demonstrate several failures to understand important principles of the science. Roberts begin by mischaracterizing "confounding", the definition of which is perhaps the main shibboleth for someone's scientist-level understanding of epidemiology. They proceed to report the myth, common to people looking for a legalistic recipe to oversimplify what is really a very complicated science, that case reports can only be used to generate scientific hypotheses and that there is something magical about peer reviewed journal articles, points that have already been addressed. Roberts improve on Colby et al. by correctly describing the cohort study and by claiming that case-control and cohort studies are just the "most common" types of epidemiology study, rather than implying that they are the only types, but they still ignore the other study types (esp. case-crossover) that are particularly applicable in the present case. They go on to report a few particular common methods as if they are the only possible methods (e.g., claiming that cohort studies only ever calculate a relative risk, while other measures like risk differences are often more useful; moreover, they describe the method for calculating only one particular relative risk measure, the risk ratio, as if it is the only choice when another measure, the odds ratio, is also quite common).

Several points like this appear, which might seem arcane to the average reader and might be necessary oversimplifications when talking to a news reporter, but seem difficult to defend in a formal report since accurate descriptions would not be much longer or more difficult. They get a few subtle and tricky points right, which suggests that they really understand some nuances of the science better than they pretend (though some of their errors seem to be based on genuine misunderstanding). This suggests that their mission is to mislead the reader into thinking epidemiology involves simple recipes and excludes more complicated reasoning, so that they can claim (when their chosen simple recipe is applied) that there is no evidence in the present case.

Again, when someone presumes to make up their own rules – and does not expect that their claims about the rules will be met with anyone who can identify their flaws – they certainly improve their chances of winning the argument.

Roberts continue to recite some overly simplistic common claims that are typically invoked by those who wish to deny most evidence. They repeat the claim that "a causal association can only be establish by the evaluation of well designed and executed epidemiologic studies", which sounds good, but the above example about the car crash and head trauma shows that it is often not true. There is no such simplistic rule. Roberts then go on to invoke the "Hill criteria", a classic piece of thinking in epidemiology but one that does not provide the simplistic rules that are typically claimed, and that modern epidemiology (see, e.g., the leading textbook in the field, Rothman and Greenland, or several of my papers) points out as being a problematic way to think about causal inference. (However, since this just seems to be company boilerplate that they cut and pasted into their report, and they make no attempt to link it to their arguments, I will not bother to challenge it in detail.)

It is difficult to believe that Roberts actually believe what they write about the journal peer review process; anyone who has worked in the system knows that it is quite often biased and politicized. They claim that peer review as been the standard since 1665; actually it is primarily a mid-to-late-20th century phenomenon (indeed, because it became a common practice only so late in his life, Einstein famously objected to a journal sending one of his submissions out for peer review, insisting that doing so was a breach of trust). They go on to describe one of the many ways in which peer review takes place, declaring it to be the way that it always works, and present an almost childlike idealization of the process. They have apparently never thought seriously about the process they are opining about; in addition to their errors they do not seem to be aware of the aforementioned point, that peer review basically never actually vouches for the accuracy of the numbers reported in an article (see Heavner et al. 2009 for an explanation of this fairly straightforward point). Nor do they seem aware that in highly politicized arenas like this one, getting the "right" answer – as defined by those who control the discourse – is critical to determining what gets the imprimatur, peer reviewed. Clearly they are writing to try to convince lawyers who have never studied the peer review process about a magical system that does not really exist. It is not clear from reading this report whether they are trying to mislead or are simply ignorant of the reality. Thus, even apart from the specific points about peer review mentioned in the substantive analysis above, these authors' apparent understanding of the process calls into question their understanding of the scientific literature.

This is quite critical, since their approach seems to be entirely premised on a misunderstanding of what constitutes useful scientific literature, and a naive preference for anything that appears in a journal. While they employ a roughly valid method for writing a review of what the literature in journals says, their further conclusion – that what those journals contain is all that science knows – it practiced naivety. If they had concluded simply "we searched some of the available evidence and if one were to consider only that evidence, it would not be sufficient to be sure there is a problem", that would have been valid. (Note the "roughly", however: They do not

explain their methods for excluding some studies they found and apparently never test their search strategy. In particular, it is notable that they do not include Pedersen, van den Berg, et al. (2009), even though the article came out several months before their report. Anyone trying to do a serious review of the literature, rather than just looking for excuses to say there is no evidence, would have known that this research had been done and anticipated the article for inclusion – after all the work of this research group represents a substantial fraction of the highly relevant literature that exists, so there is no excuse for ignoring any of it because of an arbitrary cutoff date or because it does not show up based following an arbitrary, untested search strategy.)

Someone using the Roberts methodology could never legitimately say "there is not sufficient evidence to be sure there is a problem" since they do not attempt to consider most of the evidence, which in the present case exists primarily outside of journals, a fact that Roberts no doubt knew. Moreover, if they evidence that they did not eliminate were really all that existed, the strongest conclusion that could be drawn was "we do not have enough evidence to either establish or rule out that there is a substantial problem". But since this subset of the evidence actually does point to their being some problems, the conclusion should actually be, "to the extent that we can learn much from this body of evidence in isolation, it seems reasonable to conclude that there is some problem" – roughly the same conclusion one should reach when reviewing all the available data.

Roberts' efforts to oversimplify and limit what they recognize as evidence is a reasonable tactic for the defense in a liability situation, wherein the argument "the scientific literature does not clearly establish that X causes Y" is often considered a sufficient defense, and it may be that Roberts were primarily tasked with laying down a liability defense for future tort actions. But making optimal policy requires different epistemology than does making liability awards, and they seem to be trying to obscure this. There are points to quibble with in Roberts' simplistic overview of public health, but since this appears to just be boilerplate that has little to do with the present case, those do not seem terribly important. What is more important is their section on the precautionary principle, which legitimately criticizes extremist interpretations of it, but then tries to imply that since extreme precaution is inappropriate, policy decisions should ignore all suspected health problems until (what they assert as) a burden of proof is met. Tellingly, they did not propose research that would resolve the fundamental questions, but only the research that might support claims like "those claiming health problems are lying for political reasons". After limiting scientific evidence to what they characterize as the most definitive, they declare that there was nothing "demonstrating a link between wind turbines and negative health effects" (even as their report, in multiple places, acknowledges that there are such "links"). But they carefully avoid saying that the studies they reviewed suggest that turbines do not cause substantial health risks, nor even that the available research fails to strongly suggest there is a causal relationship; the exact cautious wording, that a subset of the literature merely does not "demonstrate" it, is rather telling.

<u>Potential for gathering more information</u>

The siting of wind turbine facilities is not a situation where we cannot assess the health effects until we allow exposures to continue to be created, as might be the case with a question like "will novel environmental exposure X cause cancer". The main effects currently at issue have short term manifestations (i.e., we do not have to wait decades for cancers to develop) and there have been sufficient exposures already that information from them could be collected more systematically. The only reason we do not have better information than we do is because no one with the resources to fund the useful studies has done so. Moreover, the needed resources are relatively modest compared to what is devoted other health risks or to what is spent on building turbine facilities. In addition, it is relatively straightforward to sketch the further study that could be conducted and that would provide better information than now exists:

(a) To assess the prevalence of susceptibility (i.e., what portion of exposed people suffer disease) and similar statistics, it will be necessary to systematically collect exposed subjects for study rather than reporting only on those that volunteer or reported health problems. That is, it is necessary to find and query everyone who was exposed (residents living near turbines for some turbine facilities) or a random sample of them if the available population is too large. Subjects who could not be studied would have to be treated as missing data rather than simply left out of the study. This will provide the fraction of those exposed who experienced health problems (and would identify the distribution of specific health problems also). To provide a "denominator" to compare this to the average level of these health problems and thereby estimate how many of the cases were just background coincidence, a non-exposed group would need to be studied. (Population average statistics might be sufficient as this comparison group, but some complications about the period being studied and such might make it easier to just collect individuals for comparison.)

(b) Outcome information should be collected systematically (based on ex ante hypotheses) to avoid the risk of finding whatever diseases happen to exist and reporting them as if they were always of interest. This is critical to avoid the potential bias that results from the fact that, for any population with a particular exposure, it is almost certain that *some* set of disease can be found at an elevated level by random realizations alone. The potential for this is the most important legitimate question about the existing evidence, though this does not seem to be understood by some of those who are attacking the evidence. It is theoretically possible that the data collection methods from previous studies were designed as the data was collected, based on what was found, which inevitably creates bias and can easily create an apparent finding where no phenomenon actually exists. Fortunately there is sufficient data already to determine what diseases are likely to occur and thus what hypotheses to consider and data collection methods to design ex ante. To minimize the risk that important results will be dismissed, collected data should include as many non-subjective measures as possible, and perhaps attempt to make some non-subjective measures via instruments other than self-reporting.

(c) Systematic comparison of health outcomes in populations involuntarily that are exposed to other noxious facilities (and thus suffering the general unhappiness, but not the specific physical impacts of the turbines) would help separate diseases caused by the physical exposure from those

caused by general distress. That would help address the claim that many of the health problems result from the general distress of having been rolled over by powerful forces (a characterization that is quite close to the explanation proposed by the industry's own consultants). This would not mean that the some of the health problems are more "real" than others – suffering is still suffering – but it would suggest what portion of them could perhaps be alleviated by means other than just reducing the exposure.

(d) Relevant econometric data would provide a great deal of useful information. Real resource costs suffered by residents to respond to effects of the turbines (retrofitting houses, moving, etc.) are a good quantification of the magnitude of costs that the health effects exceeded (and thus justified the costs). Changes in local real estate values, controlling for other factors, would provide an estimate of the total perceived lowering of quality of life of living in an area as a result of the facility siting. The science cannot easily determine what portion of this is from health effects and what portion is from aesthetic and other impacts but, again, the policy viewpoint might not need to distinguish these: It is not clear that anyone would feel comfortable arguing "yes, these facilities dramatically lower property values, indicating that they do a lot of harm to local residents, but that is just because they are ugly and otherwise bothersome, not because of any formally defined disease, so the cost does not matter."

(e) A study that was designed from the top-down could model how to combine the health and economic analyses and use modern causal modeling methods. In particular, this would allow researchers to proactively properly address potential confounders rather than either ignoring them or using them as an excuse to dismiss the evidence (the frequent claim "aha, I have though of a potential confounder so therefore the conclusions are unsupported").

(f) A new study could collect detailed formal crossover data for individuals who are subject to varying levels of exposure from local turbines due to time of day, weather, and season, as well as physically leaving turbine-proximate residences. This could be combined with physical measurements of exposure, weather data, etc. Existing case study data is probably sufficient to form hypotheses about how disease outcomes change given different periods and types of exposure. People who leave the area entirely could also be followed to see what symptoms reverse, and after how long. It might also be possible to compare the effects of different turbine design details, though the study would have to be designed to assess this specifically.

(g) It would be possible to look for biomarkers of the exposure or health outcomes (e.g., stress hormones) which might help us understand the disease process and how to reduce it.

The crossover analysis would be particularly important for detecting different effects of different levels of exposure. It may not be possible to gather enough data so that comparisons between individuals will allow analysis of how effects vary by exposure detail (characteristics of the person, sound intensity, sound pitch, distance, height, weather, geology, etc.). But it may be possible to learn some of this with good crossover data.

A case for gathering more information before continuing new siting

Based on current standards of public health policy, it would clearly be appropriate to conduct most of the above studies before continuing to site facilities near residences. The studies could be completed fairly quickly based on exposures that already exist, and thus would not create indefinite delay. They could answer many of the questions that are being argued in political fights over siting, but are argued only based on limited observations and competing assertions (as well as obvious confusion about what the evidence shows), rather than the much more complete scientific analysis we could have.

The evidence might show that the magnitude of the health costs is great enough that turbines should only be sited further away from residences than current minimums, or possibly only in quite distant places (on ridgelines, offshore, or on large residence-free tracts of land) or, as explored by Pedersen et al. (2010), in areas that already have more background noise. A similar result might argue for better technology to reduce the exposure, particularly if more could be learned about exactly what characteristics of the exposure seem to cause the health problems. If it is hypothesized that particular facility designs create more problems, and thus technological changes are in order, this could be directly investigated; presumably the industry would have the incentive to point researchers toward comparing the effects of designs they think could be used with lower health effects. Other possible study results would suggest that the problem is serious but limited enough that it is most efficient to proceed but for the facility owners to openly guarantee fair compensation to local residents to compensate residents who suffer major problems. Alternatively, the research might support the claims that all health problems caused by turbines are either quite minor or quite rare, a possibility that is consistent with the existing evidence (however, note the contrast between acknowledging there is a problem but discovering it is quite rare, and thus deciding to proceed with current policies despite the costs versus declaring that there are no important problems and no costs that need to be considered even though the evidence tends to contradict this claim).

This is not a case of demanding that some vague "precautionary principle" be invoked, as in the case of banning useful plastics based on no evidence of any human health effects. The statement is not "we are ignorant and merely hypothesize that something might be wrong, so stop all action until whatever long-term studies are necessary to prove perfect safety" but rather "we have evidence that there are some health problems and we could better understand them fairly rapidly, and since they might be great enough to affect siting decisions, we should do the studies before pushing forward." To put the level of information we have and could get in perspective, there seems to be far more compelling evidence that wind turbines are causing serious health problems than there is evidence that plastics containing bisphenol-A (BPA) or electronic cigarettes (a low-risk substitute for smoking) cause any risks, but U.S. governments are moving to ban both of these despite their huge known benefits. If turbines were a pharmaceutical that was as economically beneficial as turbines are, but there was similar evidence about it causing disease side-effects, it would almost certainly not be approved without more complete study, and there is no way the industry would be able to just say "everything is really fine; please avoid doing any further study."

Our current state of relative ignorance is really more a matter of choice, since further information could be gleaned fairly easily. It is easy to understand why, under some circumstances, further research is not worth doing (Phillips 2001), perhaps because there is no reason to suspect a problem, the research would cost a fortune, or the useful research cannot be done yet. But none of those circumstances apply in this case. Arguing "we do not know enough to be sure we should change anything" is possible – right or wrong, at least someone could make the case. But it is difficult to identify any legitimate basis for arguing "we should avoid learning any more" in a case like this.

Conclusions

In summary, there is substantial evidence to support the hypothesis that wind turbines have important health effects on local residents. If forced to draw a conclusion based on existing evidence alone, it would seem defensible to conclude that there is a problem. It would certainly make little sense to conclude that there is definitely no problem, and those who make this claim offer arguments that are fundamentally unscientific. But there is simply no reason to draw a conclusion based on existing evidence alone; it is quite possible to quickly gather much more useful information than we have.

I admit to being new to this controversy and my studies have been on the content and quality of the reported science, and so there may be something hidden or political that escapes me. I have witnessed other researchers naively wandering into fields I have studied for many years, and being tricked into believing the political propaganda rather than the science. Thus I am aware of the potential limitations of understanding when someone is new to a subject matter. But as someone who specializes in trying to sort out competing epidemiology-related policy claims, I find it difficult to see how the evidence could fail to be adequate to suggest that there is a serious problem worthy of further study. The only apparent scenario that would lead to a different conclusion would be if much of the reported evidence of health problems were basically manufactured (subjects or researchers were overtly lying, or subjects were so intent on being negative that talked themselves into having diseases). But since such a scenario could only be established with further research, so even such a story leaves it impossible to justify the call to avoid further research, other than for the most cynical of motives: trying to suppress unwanted discoveries.

References

Castelo Branco NAA, Alves-Pereira M. Vibroacoustic disease. Noise & Health, 6(23):3-20, 2004.

Colby WD et al. Wind Turbine Sound and Health Effects; An Expert Panel Review. Prepared for: American Wind Energy Association and Canadian Wind Energy Association, December 2009.

Harry, A. (2007). Wind turbines, noise, and health. February 2007. http://www.wind-watch.org/documents/wp-content/uploads/wtnoise_health_2007_a_harry.pdf

Heavner K, Phillips CV, Rodu B. Peer review in epidemiology cannot accomplish its ostensible goals due to incomplete reporting and unverifiable analyses. Proceedings of the International Symposium on Peer Reviewing, at the 3rd International Conference on Knowledge Generation, Communication and Management, 2009. http://www.iiis.org/CDs2008/CD2009SCI/ispr2009/PapersPdf/V524QH.pdf

Maclure M. The Case-Crossover Design: A Method for Studying Transient Effects on the Risk of Acute Events. American Journal of Epidemiology 133(2):144-153, 1991.

Pedersen E, Waye KP. Perception and annoyance due to wind turbine noise; a dose-response relationship. Journal of the Acoustical Society of America. 116(6):3460-3470, 2004.

Pedersen E, van den Berg F, Bakker R, Bouma J. Response to noise from modern wind farms in The Netherlands. J Acoust Soc Am, 126(2):634-43, 2009.

Pedersen E, Persson WK. Wind turbine noise, annoyance and self-reported health and well-being in different living environments. Occup Environ Med , 64(7):480-6, 2007.

Pedersen E, van den Berg F, Bakker R, Bouma J. Can road traffic mask sound from wind turbines? Response to wind turbine sound at different levels of road traffic sound. Energy Policy 38(5):2520-2527, 2010.

Phillips CV. The economics of ‘more research is needed’. International Journal of Epidemiology 2001;30:771-776.

Phipps R. In the Matter of Moturimu Wind Farm Application. Evidence to the Joint Commissioners, Palmerston North. March 2007. http://www.ohariupreservationsociety.org.nz/phipps-moturimutestimony.pdf

Pierpont N *Wind turbine syndrome: a report on a natural experiment*. K-Selected Books. 2009.