


Plausibility But Not Science Has Dominated Public Discussions of the Covid Pandemic

 brownstone.org/articles/plausibility-but-not-science-has-dominated-public-discussions-of-the-covid-pandemic/

November 26, 2022



By  [Harvey Risch](#) November 26, 2022 November 26, 2022 [Philosophy](#), [Public Health](#)
30 minute read

“Attacks on me, quite frankly, are attacks on *science*.” ~ Anthony Fauci, June 9, 2021 (MSNBC).

Preposterous.

For one thing, Dr. Fauci has not reported accurately on scientific questions throughout the Covid-19 pandemic. For another, the essential dialectic of science is arguing, questioning, debating. Without debate, science is nothing more than propaganda.

Yet, one may ask, how has it been possible to present technical material to the American public, if not to the international public, for almost three years and achieve a general understanding that the matters were “scientific,” when in fact they were not? I assert that what has been fed to these publics through the traditional media over the course of the

pandemic has largely been plausibility, but not science, and that both the American and international publics, as well as most doctors, and scientists themselves, cannot tell the difference. However, the difference is fundamental and profound.

Science starts with theories, hypotheses, that have examinable empiric ramifications. Nevertheless, those theories are not science; they *motivate* science. Science occurs when individuals do experiments or make observations that bear upon the implications or ramifications of the theories. Those findings tend to support or refute the theories, which are then modified or updated to adjust to the new observations or discarded if compelling evidence shows that they fail to describe nature. The cycle is then repeated. *Science* is the performance of empirical or observational work to obtain evidence confirming or refuting theories.

In general, theories tend to be plausible statements describing something specific about how nature operates. Plausibility is in the eye of the beholder, since what is plausible to a technically knowledgeable expert may not be plausible to a lay person. For example—perhaps oversimplified—heliocentrism was not plausible before Nicolaus Copernicus published his theory in 1543, and it was not particularly plausible afterward for quite some time, until Johannes Kepler understood that astronomical measurements made by Tycho Brahe suggested refining the Copernican circular orbits to ellipses, as well as that mathematical rules seemed to govern the planetary motions along those ellipses—yet reasons for those mathematical rules, even if they were good descriptions of the motions, weren't plausible until Isaac Newton in 1687 posited the existence of a universal gravitational force between masses, along with a mass-proportional, inverse-square distance law governing the magnitude of the gravitational attraction, and observed numerous quantitative phenomena consistent with and supporting this theory.

For us today, we hardly think about the plausibility of elliptic heliocentric solar system orbits, because observational data spanning 335 years have been highly consistent with that theory. But we might balk at thinking it plausible that light travels simultaneously as both particles and waves, and that making measurements on the light, what we do as observers, determines whether we see particle behavior or wave behavior, and we can choose to observe either particles or waves, but not both at the same time. Nature is not necessarily plausible.

But all the same, plausible theories are easy to believe, and that is the problem. That is what we have been fed for almost three years of the Covid-19 pandemic. In fact though, we have been fed plausibility instead of science for much longer.

Cargo-Cult Science

Charlatans purporting to bend spoons with their minds, or claiming to study unconfirmable, irreplicable “extrasensory perception” were very popular in the 1960s and 1970s. Strange beliefs in what “science” could establish reached such a level that physics Nobel Laureate Richard Feynman delivered the 1974 Caltech commencement address (Feynman, 1974) bemoaning such irrational beliefs. His remarks were not aimed at the general public, but at graduating Caltech students, many of whom were destined to become academic scientists.

In his address, Feynman described how South Sea Islanders, after World War II, mimicked US soldiers stationed there during the war who had guided airplane landings of supplies. The island residents, using local materials, reproduced the form and behaviors of what they had witnessed of the American GIs, but no supplies came.

In our context, Feynman’s point would be that until a theory has objective empirical evidence bearing upon it, it remains only a theory no matter how plausible it may seem to everyone who entertains it. The Islanders were missing the crucial fact that they did not understand how the supply system worked, in spite of how plausible their reproduction of it was to them. That Feynman felt compelled to warn graduating Caltech students of the difference between plausibility and science, suggesting that this difference was not adequately learned in their Institute educations. It was not explicitly taught when this author was an undergraduate there in those years, but somehow, we were expected to have learned it “by osmosis.”

Evidence-Based Medicine

There is perhaps no bigger plausibility sham today than “evidence-based medicine” (EBM). This term was coined by Gordon Guyatt in 1990, after his first attempt, “Scientific Medicine,” failed to gain acceptance the previous year. As a university epidemiologist in 1991, I was insulted by the hubris and ignorance in the use of this term, EBM, as if medical evidence were somehow “unscientific” until proclaimed a new discipline with new rules for evidence. I was not alone in criticism of EBM (Sackett et al., 1996), though much of that negative response seems to have been based on loss of narrative control rather than on objective review of what medical research had actually accomplished without “EBM.”

Western medical knowledge has accreted for thousands of years. In the Hebrew Bible (Exodus 21:19), “When two parties quarrel and one strikes the other ... the victim shall be made thoroughly healed” [my translation] which implies that individuals who had types of medical knowledge existed and that some degree of efficacy inhered. Hippocrates, in the fifth-fourth century BCE, suggested that disease development might not be random but related to exposures from the environment or to certain behaviors. In that era, there were plenty of what today we would consider counterexamples to good medical practice. Nevertheless, it was a start, to think about rational evidence for medical knowledge.

James Lind (1716-1794) advocated for scurvy protection through the eating of citrus. This treatment was known to the ancients, and in particular had been earlier recommended by the English military surgeon John Woodall (1570-1643)—but Woodall was ignored. Lind gets the credit because in 1747 he carried out a small but successful nonrandomized, controlled trial of oranges and lemons vs other substances among 12 scurvy patients.

During the 1800s, Edward Jenner's use of cowpox as a smallpox vaccine was elaborated by culturing in other animals and put into general use in outbreaks, so that by the time of the 1905 Supreme Court case of *Jacobson v. Massachusetts*, the Chief Justice could assert that smallpox vaccination was agreed upon by medical authorities to be a commonly accepted procedure. Medical journals started regular publications also in the 1800s. For example, the *Lancet* began publishing in 1824. Accreting medical knowledge started to be shared and debated more generally and widely.

Fast-forward to the 1900s. In 1914-15, Joseph Goldberger (1915) carried out a nonrandomized dietary intervention trial that concluded that pellagra was caused by lack of dietary protein. In the 1920s, vaccines for diphtheria, pertussis, tuberculosis and tetanus were developed. Insulin was extracted. Vitamins, including Vitamin D for preventing rickets, were developed. In the 1930s, antibiotics began to be created and used effectively. In the 1940s, acetaminophen was developed, as were chemotherapies, and conjugated estrogen began to be used to treat menopausal hot flashes. Effective new medications, vaccines and medical devices grew exponentially in number in the 1950s and 1960s. All without EBM.

In 1996, responding to criticisms of EBM, David Sackett et al. (1996) attempted to explain its overall principles. Sackett asserted that EBM followed from "Good doctors use both individual clinical expertise and the best available external evidence." This is an anodyne plausibility implication, but both components are basically wrong or at least misleading. By phrasing this definition in terms of what individual doctors should do, Sackett was implying that individual practitioners should use their own clinical observations and experience. However, the general evidential representativeness of one individual's clinical experience is likely to be weak. Just like other forms of evidence, clinical evidence needs to be systematically collected, reviewed, and analyzed, to form a synthesis of clinical reasoning, which would then provide the clinical component of scientific medical evidence.

A bigger failure of evidential reasoning is Sackett's statement that one should use "the best available external evidence" rather than *all* valid external evidence. Judgments about what constitutes "best" evidence are highly subjective and do not necessarily yield overall results that are quantitatively the most accurate and precise (Hartling et al., 2013; Bae, 2016). In formulating his now canonical "aspects" of evidential causal reasoning, Sir Austin Bradford Hill (1965) did not include an aspect of what would constitute "best" evidence, nor did he suggest that studies should be measured or categorized for "quality of study" nor even that some types of study designs might be intrinsically better than others. In the *Reference Manual on Scientific Evidence*, Margaret Berger (2011) states explicitly, "... many of the most

well-respected and prestigious scientific bodies (such as the International Agency for Research on Cancer (IARC), the Institute of Medicine, the National Research Council, and the National Institute for Environmental Health Sciences) consider all the relevant available scientific evidence, taken as a whole, to determine which conclusion or hypothesis regarding a causal claim is best supported by the body of evidence.” This is exactly Hill’s approach; his aspects of causal reasoning have been very widely used for more than 50 years to reason from observation to causation, both in science and in law. That EBM is premised on subjectively cherry-picking “best” evidence is a plausible method but not a scientific one.

Over time, the EBM approach to selectively considering “best” evidence seems to have been “dumbed down,” first by placing randomized controlled trials (RCTs) at the top of a pyramid of all study designs as the supposed “gold standard” design, and later, as the asserted only type of study that can be trusted to obtain unbiased estimates of effects. All other forms of empirical evidence are “potentially biased” and therefore unreliable. This is a plausibility conceit as I will show below.

But it is so plausible that it is routinely taught in modern medical education, so that most doctors only consider RCT evidence and dismiss all other forms of empirical evidence. It is so plausible that this author had an on-air verbal battle over it with a medically uneducated television commentator who provided no evidence other than plausibility (Whelan, 2020): Isn’t it “just obvious” that if you randomize subjects, any differences must be caused by the treatment, and no other types of studies can be trusted? Obvious, yes; true, no.

Who benefits from a sole, obsessive focus on RCT evidence? RCTs are very expensive to conduct if they are to be epidemiologically valid and statistically adequate. They can cost millions or tens of millions of dollars, which limit their appeal largely to companies promoting medical products likely to bring in profits substantially larger than those costs. Historically, pharma control and manipulation of RCT evidence in the regulation process provided an enormous boost in the ability to push products through regulatory approval into the marketplace, and the motivation to do this still continues today.

This problem was recognized by Congress, which passed the Food and Drug Administration Modernization Act of 1997 (FDAMA) that established in 2000 the [ClinicalTrials.gov](https://www.clinicaltrials.gov) website for registration of all clinical trials performed under investigational new drug applications to examine the effectiveness of experimental drugs for patients with serious or life-threatening conditions (National Library of Medicine, 2021). For related reasons involving conflicts of interests in clinical trials, the ProPublica “Dollars for Docs” website (Tigas et al., 2019) covering pharma company payments to doctors over the years 2009-2018 and the OpenPayments website (Centers for Medicare & Medicaid Services, 2022) covering payments from 2013 through 2021 were established and made publicly searchable. These information systems were created because the “plausibility” that randomization automatically makes study results accurate and unbiased was recognized as insufficient to cope with research chicanery and inappropriate investigator conflict-of-interest motives.

While these attempts to reform or limit medical research corruption have helped, misrepresentation of evidence under the guise of EBM persists. One of the worst examples was a paper published in the *New England Journal of Medicine* February 13, 2020, at the beginning of the Covid-19 pandemic, titled, “The Magic of Randomization versus the Myth of Real-World Evidence,” by four well-known British medical statisticians having substantial ties to pharma companies (Collins et al., 2020). It was likely written in January 2020, before most people knew that the pandemic was coming. This paper claims that randomization automatically creates strong studies, and that all nonrandomized studies are evidentiary rubbish. At the time of reading it, I felt it to be a screed against my entire discipline, epidemiology. I was immediately offended by it, but I later understood the serious conflicts of interest of the authors. Representing that only highly unaffordable RCT evidence is appropriate for regulatory approvals provides a tool for pharma companies to protect their expensive, highly profitable patent products against competition by effective and inexpensive off-label approved generic medications whose manufacturers would not be able to afford large-scale RCTs.

Randomization

So, what is the flaw of randomization to which I have been alluding, that requires a deeper examination in order to understand the relative validity of RCT studies vs other study designs? The problem lies in the understanding of *confounding*. Confounding is an epidemiological circumstance where a relationship between an exposure and an outcome is not due to the exposure, but to a third factor (the confounder), at least in part. The confounder is somehow associated with the exposure but is not a result of the exposure.

In such cases, the apparent exposure-outcome relationship is really due to the confounder-outcome relationship. For example, a study of alcohol consumption and cancer risk could be potentially confounded by smoking history which correlates with alcohol use (and isn't caused by alcohol use) but is really driving the increased cancer risk. A simple analysis of alcohol and cancer risk, ignoring smoking, would show a relationship. However, once the effect of smoking was controlled or adjusted, the alcohol relationship with cancer risk would decline or disappear.

The purpose of randomization, of balancing everything between the treatment and control groups, is to remove potential confounding. Is there any other way to remove potential confounding? Yes: measure the factors in question and adjust or control for them in statistical analyses. It is thus apparent that randomization has exactly one possible benefit not available to nonrandomized studies: the control of *unmeasured* confounders. If biological, medical, or epidemiological relationships are incompletely understood about an outcome of interest, then not all relevant factors may be measured, and some of those unmeasured factors could still confound an association of interest.

Thus, randomization, *in theory*, removes potential confounding by unmeasured factors as an explanation for an observed association. That is the plausibility argument. The question though concerns how well randomization works in reality, and who exactly needs to be balanced by the randomization. Clinical trials apply randomization to all participating subjects to determine treatment group assignments. If in the study outcome event individuals comprise a subset of the total study, then those outcome people need to be balanced in their potential confounders as well. For example, if all of the deaths in the treatment group are males and all in the placebo group are females, then gender likely confounds the effect of treatment.

The problem is, RCT studies essentially never explicitly demonstrate adequate randomization of their outcome subjects, and what they purport to show of randomization for their total treatment groups is almost always scientifically irrelevant. This problem likely arises because the individuals carrying out RCT studies, and the reviewers and journal editors who consider their papers, do not sufficiently understand epidemiologic principles.

In most RCT publications, the investigators provide a perfunctory initial descriptive table of the treatment and placebo groups (as columns), vs various measured factors (as rows). That is, the percent distributions of treatment and placebo subjects by gender, age group, race/ethnicity etc. The third column in these tables is usually the p-value statistic for the frequency difference between the treatment and placebo subjects on each measured factor. Loosely speaking, this statistic estimates a probability that a frequency difference between treatment and placebo subjects this large could have occurred by chance. Given that the subjects were assigned their treatment groups entirely by chance, statistical examination of the randomization chance process is tautological and irrelevant. That in some RCTs, some factors may appear to be more extreme than chance would allow under randomization is only because multiple factors down the rows have been examined for distributional differences and in such circumstances, statistical control of multiple comparisons must be invoked.

What is needed in the third column of the RCT descriptive table is not p-value, but a measure of the magnitude of confounding of the particular row factor. *Confounding is not measured by how it occurred, but by how bad it is.* In my experience as a career epidemiologist, the best single measure of confounding is the percentage change in the magnitude of the treatment-outcome relationship with vs without adjustment for the confounder. So for example, if with adjustment for gender, treatment cuts mortality by 25% (relative risk = 0.75), but without adjustment cuts it by 50%, then the magnitude of confounding by gender would be $(0.75 - 0.50)/0.50 = 50\%$. Epidemiologists generally consider more than a 10% change with such adjustment to imply that confounding is present and needs to be controlled.

As I have observed, most RCT publications do not provide the magnitude of confounding estimates for their overall treatment groups, and never for their outcome subjects. So it is not possible to tell that the outcome subjects have been adequately randomized for all of the factors given in the paper's descriptive table. But the potential fatal flaw of RCT studies, what can make them no better than nonrandomized studies and in some cases worse, is that randomization only works when large numbers of subjects have been randomized (Deaton and Cartwright, 2018), and this applies specifically to the outcome subjects, not just to the total study.

Consider flipping a coin ten times. It might come up at least seven heads and three tails, or vice versa, easily by chance (34%). However, the magnitude of this difference, $7/3 = 2.33$, is potentially quite large in terms of possible confounding. On the other hand, occurrence of the same 2.33 magnitude from 70 or more heads out of 100 flips would be rare, $p=.000078$. In order for randomization to work, there needs to be sizable numbers of outcome events in both the treatment and placebo groups, say 50 or more in each group. This is the unspoken potential major flaw of RCT studies that makes their plausibility argument useless, because RCT studies are generally designed to have enough statistical power to find statistical significance of their primary result if the treatment works as predicted, but not designed to have enough outcome subjects to reduce potential confounding to less than 10% say.

An important example of this issue can be seen in the first published efficacy RCT result for the Pfizer BNT162b2 mRNA Covid-19 vaccine (Polack et al., 2020). This study was considered large enough (43,548 randomized participants) and important enough (Covid-19) that because of its assumed RCT plausibility it secured publication in the "prestigious" *New England Journal of Medicine*. The primary outcome of the study was the occurrence of Covid-19 with onset at least seven days after the second dose of the vaccine or placebo injection. However, while it observed 162 cases among the placebo subjects, enough for good randomization, it found only eight cases among the vaccine subjects, nowhere nearly enough for randomization to have done anything to control confounding.

From general epidemiologic experience, an estimated relative risk this large (approximately $162/8 = 20$) would be unlikely entirely to be due to confounding, but the accuracy of the relative risk or its implied effectiveness ($(20 - 1)/20 = 95\%$) is in doubt. That this vaccine in use was observed not to be this effective in reducing infection risk is not surprising given the weakness of the study result because of inadequate sample size to assure that randomization worked for the outcome subjects in both the treatment and placebo groups.

This "dive into the weeds" of epidemiology illuminates why an RCT study with fewer than, say, 50 outcome subjects in each and every treatment arm of the trial has little to no claim to avoiding possible confounding by unmeasured factors. But it also makes evident why such a trial may be *worse* than a nonrandomized controlled trial of the same exposure and outcome.

In nonrandomized trials, the investigators know that many factors may, as possible confounders, influence the occurrence of the outcome, so they measure everything they think relevant, in order to then adjust and control for those factors in the statistical analyses.

However, in RCTs, investigators routinely think that the randomization has been successful and thus carry out unadjusted statistical analyses, providing potentially confounded results. When you see RCTs paraded as “large” studies because of their tens of thousands of participants, look past that, to the numbers of primary outcome events in the treatment arms of the trial. Trials with small numbers of primary outcome events are useless and should not be published, let alone relied upon for public health or policy considerations.

Empirical Evidence

After reading all of the foregoing, you might think that these arguments concerning randomized vs nonrandomized trials are very plausible, but what about empirical evidence to support them? For that, a very thorough analysis was carried out by the Cochrane Library Database of Systematic Reviews (Anglemyer et al., 2014). This study comprehensively searched seven electronic publication databases for the period from January 1990 through December 2013, to identify all systematic review papers that compared “quantitative effect size estimates measuring efficacy or effectiveness of interventions tested in [randomized] trials with those tested in observational studies.” In effect a meta-analysis of meta-analyses, the analysis included many thousands of individual study comparisons as summarized across 14 review papers.

The bottom line: an average of only 8% difference (95% confidence limits, -4% to 22%, not statistically significant) between the RCTs and their corresponding nonrandomized trials results. In summary, this body of knowledge—the empirical as well as that based upon epidemiologic principles—demonstrates that, contra so-called “plausibility,” randomized trials have no automatic ranking as a gold standard of medical evidence or as the only acceptable form of medical evidence, and that every study needs to be critically and objectively examined for its own strengths and weaknesses, and for how much those strengths and weaknesses matter to the conclusions drawn.

Other Plausibilities

During the Covid-19 pandemic, numerous other assertions of scientific evidence have been used to justify public health policies, including for the very declaration of the pandemic emergency itself. Underlying many of these has been the plausible but fallacious principle that the goal of public health pandemic management is to minimize the number of people infected by the SARS-CoV-2 virus.

That policy may seem obvious, but it is wrong as a blanket policy. What needs to be minimized are the harmful consequences of the pandemic. If infection leads to unpleasant or annoying symptoms for most people but no serious or long-term issues—as is generally the case with SARS-CoV-2, particularly in the Omicron era—then there would be no tangible benefit of general public-health interventions and limitations infringing upon natural or economic rights of such individuals and causing harms in themselves.

Western societies, including the US, take annual respiratory infection waves in stride without declared pandemic emergencies, even though they produce millions of infected individuals each year, because the consequences of infection are considered generally medically minor, even allowing for some tens of thousands of deaths annually.

It was established in the first few months of the Covid-19 pandemic that the infection mortality risk varied by more than 1,000-fold across the age span, and that people without chronic health conditions such as diabetes, obesity, heart disease, kidney disease, cancer history etc., were at negligible risk of mortality and very low risk of hospitalization. At that point, it was straightforward to define categories of high-risk individuals who on average would benefit from public health interventions, vs low-risk individuals who would successfully weather the infection without appreciable or long-term issues. Thus, an obsessive, one-size-fits-all pandemic management scheme that did not distinguish risk categories was unreasonable and oppressive from the outset.

Accordingly, measures promoted by plausibility to reduce infection transmission, even had they been effective for that purpose, have not served good pandemic management. These measures however were never justified by scientific evidence in the first place. The Six-Foot Social Distancing Rule was an arbitrary concoction of the CDC (Dangor, 2021). Claims of benefit for wearing of face masks have rarely distinguished potential benefit to the wearer—for whom such wearing would be a personal choice whether or not to accept more theoretical risk—vs benefit to bystanders, so-called “source control,” wherein public health considerations might properly apply. Studies of mask-based source control for respiratory viruses, where the studies are without fatal flaws, have shown no appreciable benefit in reducing infection transmission (Alexander, 2021; Alexander, 2022; Burns, 2022).

General population lockdowns have never been used in Western countries and have no evidence of effect for doing anything other than postponing the inevitable (Meunier, 2020), as Australia population data make clear (Worldometer, 2022). In the definitive discussion of public health measures for control of pandemic influenza ([Inglesby et al., 2006](#)), the authors state, “There are no historical observations or scientific studies that support the confinement by quarantine of groups of possibly infected people for extended periods in order to slow the spread of influenza. A World Health Organization (WHO) Writing Group, after reviewing the literature and considering contemporary international experience, concluded that ‘forced isolation and quarantine are ineffective and impractical.’ ... The negative consequences of large-scale quarantine are so extreme (forced confinement of sick people with the well;

complete restriction of movement of large populations; difficulty in getting critical supplies, medicines, and food to people inside the quarantine zone) that this mitigation measure should be eliminated from serious consideration.”

On travel restrictions, Inglesby et al. (2006) note, “Travel restrictions, such as closing airports and screening travelers at borders, have historically been ineffective. The World Health Organization Writing Group concluded that ‘screening and quarantining entering travelers at international borders did not substantially delay virus introduction in past pandemics ... and will likely be even less effective in the modern era.’” On school closures (Inglesby et al., 2006): “In previous influenza epidemics, the impact of school closings on illness rates has been mixed. A study from Israel reported a decrease in respiratory infections after a 2-week teacher strike, but the decrease was only evident for a single day. On the other hand, when schools closed for a winter holiday during the 1918 pandemic in Chicago, ‘more influenza cases developed among pupils ... than when schools were in session.’”

This discussion makes clear that these actions supposedly interfering with virus transmission on the basis of plausibility arguments for their effectiveness have been both misguided for managing the pandemic, and unsubstantiated by scientific evidence of effectiveness in reducing spread. Their large-scale promotion has demonstrated the failure of public-health policies in the Covid-19 era.

Plausibility vs Bad Science

An argument could be entertained that various public-health policies as well as information made available to the general public have not been supported by plausibility but instead by bad or fatally flawed science, posing as real science. For example, in its in-house, non-peer-reviewed journal, *Morbidity and Mortality Weekly Reports*, CDC has published a number of analyses of vaccine effectiveness. These reports described cross-sectional studies but analyzed them as if they were case-control studies, systematically using estimated odds ratio parameters instead of relative risks to calculate vaccine effectiveness. When study outcomes are infrequent, say fewer than 10% of study subjects, then odds ratios can approximate relative risks, but otherwise, odds ratios tend to be overestimates. However, in cross-sectional studies, relative risks can be directly calculated and can be adjusted for potential confounders by relative-risk regression (Wacholder, 1986), similar to the use of logistic regression in case-control studies.

A representative example is a study of the effectiveness of third-dose Covid-19 vaccines (Tenforde et al., 2022). In this study, “... the IVY Network enrolled 4,094 adults aged ≥ 18 years,” and after relevant subject exclusions, “2,952 hospitalized patients were included (1,385 case-patients and 1,567 non-COVID-19 controls).” Cross-sectional studies—by design—identify total numbers of subjects, whereas the numbers of cases and controls, and exposed and unexposed, happen outside of investigator intervention, i.e., by whatever natural processes underlie the medical, biological and epidemiological mechanisms under

examination. By selecting a total number of subjects, the Tenforde et al. study is by definition a cross-sectional design. This study reported a vaccine effectiveness of 82% among patients without immunocompromising conditions. This estimate reflects an adjusted odds ratio of $1 - 0.82 = 0.18$. However, the fraction of case patients among the vaccinated was 31% and among the unvaccinated was 70%, neither of which is sufficiently infrequent to allow use of the odds ratio approximation to calculate vaccine effectiveness. By the numbers in the study report Table 3, I calculate an unadjusted relative risk of 0.45 and an approximately adjusted relative risk of 0.43, giving the true vaccine effectiveness of $1 - 0.43 = 57\%$ which is substantially different and much worse than the 82% presented in the paper.

In a different context, after I published a summary review article on the use of hydroxychloroquine (HCQ) for early outpatient Covid-19 treatment (Risch, 2020), a number of clinical trials papers were published in an attempt to show that HCQ is ineffective. The first of these so-called “refutations” were conducted in hospitalized patients, whose disease is almost entirely different in pathophysiology and treatment than early outpatient illness (Park et al., 2020). The important outcomes that I had addressed in my review, risks of hospitalization and mortality, were distracted in these works by focus on subjective and lesser outcomes such as duration of viral test positivity, or length of hospital stay.

Subsequently, RCTs of outpatient HCQ use began to be published. A typical one is that by Caleb Skipper et al. (2020). The primary endpoint of this trial was a change in overall self-reported symptom severity over 14 days. This subjective endpoint was of little pandemic importance, especially given that the subjects in studies by this research group were moderately able to tell whether they were in the HCQ or placebo arms of the trial (Rajasingham et al., 2021) and thus the self-reported outcomes were not all that blinded to the medication arms. From their statistical analyses, the authors appropriately concluded that “Hydroxychloroquine did not substantially reduce symptom severity in outpatients with early, mild COVID-19.” However, the general media reported this study as showing that “hydroxychloroquine doesn’t work.” For example, Jen Christensen (2020) in *CNN Health* stated about this study, “The antimalarial drug hydroxychloroquine did not benefit non-hospitalized patients with mild Covid-19 symptoms who were treated early in their infection, according to a study published Thursday in the medical journal *Annals of Internal Medicine*.”

But in fact, the Skipper study did report on the two outcomes of importance, risks of hospitalization and mortality: with placebo, 10 hospitalizations and 1 death; with HCQ, 4 hospitalizations and 1 death. These numbers show a 60% reduced risk of hospitalization which, though not statistically significant ($p=0.11$), is entirely consistent with all other studies of hospitalization risk for HCQ use in outpatients (Risch, 2021). Nevertheless, these small numbers of outcome events are not nearly enough for randomization to have balanced any factors, and the study is essentially useless on this basis. But it was still misinterpreted in the lay literature as showing that HCQ provides no benefit in outpatient use.

Conclusions

Many other instances of plausible scientific claptrap or bad science have occurred during the Covid-19 pandemic. As was seen with the retracted Surgisphere papers, medical journals routinely and uncritically publish this nonsense as long as conclusions align with government policies. This body of fake knowledge has been promulgated at the highest levels, by the NSC, FDA, CDC, NIH, WHO, Wellcome Trust, AMA, medical specialty boards, state and local public health agencies, multinational pharma companies and other organizations around the world that have violated their responsibilities to the public or have purposely chosen not to understand the fake science.

The US Senate recently voted, for the third time, to end the Covid-19 state of emergency, yet President Biden stated that he would veto the measure because of “fear” of recurring case numbers. My colleagues and I argued almost a year ago that the pandemic emergency was over (Risch et al., 2022), yet the spurious reliance on case counts to justify suppression of human rights under the cover of “emergency” continues unabated.

Massive censorship by the traditional media and much of social media has blocked most public discussion of this bad and fake science. Censorship is the tool of the undefendable, since valid science inherently defends itself. Until the public begins to understand the difference between plausibility and science and how large the effort has been to mass-produce science “product” that looks like science but is not, the process will continue and leaders seeking authoritarian power will continue to rely on it for fake justification.

References

Alexander, P. E. (2021, December 20). *More than 150 Comparative Studies and Articles on Mask Ineffectiveness and Harms*. Brownstone Institute. <https://brownstone.org/articles/more-than-150-comparative-studies-and-articles-on-mask-ineffectiveness-and-harms/>

Alexander, P. E. (2022, June 3). *CDC Refuses to Post the Fix to Its Mask Study*. Brownstone Institute. <https://brownstone.org/articles/cdc-refuses-to-post-the-fix-to-its-mask-study/>

Anglemyer, A., Horvath, H. T., Bero, L. (2014). Healthcare outcomes assessed with observational study designs compared with those assessed in randomized trials (Review). *Cochrane Database of Systematic Reviews*, 4, Article MR000034. <https://doi.org/10.1002/14651858.MR000034.pub2>

Bae, J.-M. (2016). A suggestion for quality assessment in systematic reviews of observational studies in nutritional epidemiology. *Epidemiology and Health*, 38, Article e2016014. <https://doi.org/10.4178/epih.e2016014>

Berger, M. A. (2011). The admissibility of expert testimony. In National Research Council, Committee on the Development of the Third Edition of the Reference Manual on Scientific Evidence, *Reference Manual on Scientific Evidence*, Third Edition (pp. 11-36). National

Academies Press. <https://nap.nationalacademies.org/catalog/13163/reference-manual-on-scientific-evidence-third-edition>

Burns, E. (2022, November 10). *Another Day, Another Terrible Mask Study. Let's look under the hood of the newest piece of low quality science on masks.* Substack. <https://emilyburns.substack.com/p/another-day-another-terrible-mask>

Centers for Medicare & Medicaid Services. (2022, June). *Search Open Payments.* U.S. Department of Health and Human Services, Centers for Medicare & Medicaid Services. <https://openpaymentsdata.cms.gov/>

Christensen, J. (2020, July 16). *Hydroxychloroquine also doesn't help Covid-19 patients who aren't hospitalized, new study finds.* CNN Health. <https://www.cnn.com/2020/07/16/health/hydroxychloroquine-doesnt-work-hospitalized-patients/>

Collins, R., Bowman, L., Landray, M., & Peto, R. (2020). The Magic of Randomization versus the Myth of Real-World Evidence. *New England Journal of Medicine*, 382(7), 674-678. <https://www.nejm.org/doi/10.1056/NEJMs1901642>

Dangor, G. (2021, September 19). *CDC's Six-Foot Social Distancing Rule Was 'Arbitrary', Says Former FDA Commissioner.* Forbes. <https://www.forbes.com/sites/graisondangor/2021/09/19/cdcs-six-foot-social-distancing-rule-was-arbitrary-says-former-fda-commissioner/>

Deaton, A., & Cartwright, N. (2018). Understanding and misunderstanding randomized controlled trials. *Social Science & Medicine*, 210, 2-21. <https://doi.org/10.1016/j.socscimed.2017.12.005>

Feynman, R. P. (1974). Cargo Cult Science. *Engineering and Science*, 37(7), 10-13. <https://resolver.caltech.edu/CaltechES:37.7.CargoCult>

Goldberger, J., Waring, C. H., & Willets, D. G. (1915). The prevention of pellagra: A test of diet among institutional inmates. *Public Health Reports*, 30(43), 3117-3131. <https://www.jstor.org/stable/4572932>

Hartling, L., Milne, A., Hamm, M. P., Vandermeer, B., Ansari, M., Tsertsvadze, A., Dryden, D. M. (2013). Testing the Newcastle Ottawa Scale showed low reliability between individual reviewers. *Journal of Clinical Epidemiology*, 66, 982-993. <https://doi.org/10.1016/j.jclinepi.2013.03.003>

Hill, A. B. (1965). The environment and disease: association or causation. *Proceedings of the Royal Society of Medicine*, 58(5), 295-300. <https://doi.org/10.1177/003591576505800503>

Inglesby, T. V., Nuzzo, J. B., O'Toole, T., Henderson, D. A. (2006). Disease mitigation measures in the control of pandemic influenza. *Biosecurity and Bioterrorism: Biodefense Strategy, Practice, and Science*, 4(4):366-375. <https://doi.org/10.1089/bsp.2006.4.366>

Meunier, T. (2020, May 1). *Full lockdown policies in Western Europe countries have no evident impacts on the COVID-19 epidemic*. medRxiv. <https://doi.org/10.1101/2020.04.24.20078717>

MSNBC. (2021, June 9). *Fauci responds to attacks from Republicans* [Video]. YouTube. <https://www.youtube.com/watch?v=z-tfZr8lv0s>

National Library of Medicine (2021, May). *ClinicalTrials.gov. History, Policies, and Laws*. U.S. Department of Health and Human Services, National Institutes of Health, National Library of Medicine. <https://clinicaltrials.gov/ct2/about-site/history>

Park, J. J. H., Declodt, E. H., Rayner, C. R., Cotton, M., Mills, E. J. (2020). Clinical trials of disease stages in COVID 19: complicated and often misinterpreted. *Lancet Global Health*, 8(10), e1249-e1250. [https://doi.org/10.1016/S2214-109X\(20\)30365-X](https://doi.org/10.1016/S2214-109X(20)30365-X)

Polack, F. P., Thomas, S. J., Kitchin, N., Absalon, J., Gurtman, A., Lockhart, S., Perez, J. L., Pérez Marc, G., Moreira, E. D., Zerbini, C., Bailey, R., Swanson, K. A., Roychoudhury, S., Koury, K., Li, P., Kalina, W. V., Cooper, D., Frenck, R. W., Jr., Hammitt, L. L., ..., Gruber, W. C. (2020). Safety and efficacy of the BNT162b2 mRNA Covid-19 vaccine. *New England Journal of Medicine*, 383(27), 2603-2615. <https://www.nejm.org/doi/full/10.1056/nejmoa2034577>

Rajasingham, R., Bangdiwala, A. S., Nicol, M. R., Skipper, C. P., Pastick, K. A., Axelrod, M. L., Pullen, M. F., Nascene, A. A., Williams, D. A., Engen, N. W., Okafor, E. C., Rini, B. I., Mayer, I. A., McDonald, E. G., Lee, T. C., Li P., MacKenzie, L. J., Balko, J. M., Dunlop, S. J., ..., Lofgren, S. M. (2021). Hydroxychloroquine as Pre-exposure Prophylaxis for Coronavirus Disease 2019 (COVID-19) in Healthcare Workers: A Randomized Trial. *Clinical Infectious Diseases*, 72(11), e835-e843. <https://doi.org/10.1093/cid/ciaa1571>

Risch, H. A. (2020). Early Outpatient Treatment of Symptomatic, High-Risk COVID-19 Patients That Should Be Ramped Up Immediately as Key to the Pandemic Crisis. *American Journal of Epidemiology*, 189(11), 1218-1226. <https://doi.org/10.1093/aje/kwaa093>

Risch, H. A. (2021, June 17). *Hydroxychloroquine in Early Treatment of High-Risk COVID-19 Outpatients: Efficacy and Safety Evidence*. EarlyCovidCare.org, <https://earlycovidcare.org/wp-content/uploads/2021/09/Evidence-Brief-Risch-v6.pdf>

Risch, H., Bhattacharya, J., Alexander, P. E. (2022, January 23). *The Emergency Must Be Ended, Now*. Brownstone Institute. <https://brownstone.org/articles/the-emergency-must-be-ended-now/>

Sackett, D. L., Rosenberg, W. M. C., Gray, J. A. M., Haynes, R. B., & Richardson, W. S. (1996). Evidence based medicine: what it is and what it isn't. *BMJ*, 312, Article 71. <https://doi.org/10.1136/bmj.312.7023.71>

Skipper, C. P., Pastick, K. A., Engen, N. W., Bangdiwala, A. S., Abassi, M., Lofgren, S. M., Williams, D. A., Okafor, E. C., Pullen, M. F., Nicol, M. R., Nascene, A. A., Hullsiek, K. H., Cheng, M. P., Luke, D., Lothar, S. A., MacKenzie, L. J., Drobot, G., Kelly, L. E., Schwartz, I. S., ..., Boulware, D. R. (2020). Hydroxychloroquine in Nonhospitalized Adults With Early COVID-19 : A Randomized Trial. *Annals of Internal Medicine*, 173(8), 623-631. <https://doi.org/10.7326/M20-4207>

Tenforde, M. W., Patel, M. M., Gaglani, M., Ginde, A. A., Douin, D. J., Talbot, H. K., Casey, J. D., Mohr, N. M., Zepeski, A., McNeal, T., Ghamande, S., Gibbs, K. W., Files, D. C., Hager, D. N., Shehu, A., Prekker, M. E., Erickson, H. L., Gong, M. N., Mohamed, A., ..., Self, W. H. (2022). Morbidity and Mortality Weekly Report, 71(4), 118-124. <https://www.cdc.gov/mmwr/volumes/71/wr/mm7104a2.htm>

Tigas, M., Jones, R. G., Ornstein, C., & Groeger, L. (2019, October 17). *Dollars for Docs. How Industry Dollars Reached Your Doctors*. ProPublica. <https://projects.propublica.org/docdollars/>

Wacholder, S. (1986). Binomial regression in GLIM: estimating risk ratios and risk differences. *American Journal of Epidemiology*, 123(1), 174-184. <https://doi.org/10.1093/oxfordjournals.aje.a114212>

Whelan, R. (2020, August 3). *2020-08-03 – CNN COVID with Interview Harvey Risch, Yale Epidemiologist* [Video]. YouTube. <https://www.youtube.com/watch?v=gGO6Ke81bUQ>

Worldometer. (2022, November 15). *Total Coronavirus Cases in Australia*. Worldometer. <https://www.worldometers.info/coronavirus/country/australia/>

Author



Harvey Risch

Harvey Risch, Senior Scholar at Brownstone Institute, is a physician and a Professor Emeritus of Epidemiology at Yale School of Public Health and Yale School of Medicine. His main research interests are in cancer etiology, prevention and early diagnosis, and in epidemiologic methods.