Jeremy Berg 13h

Bluetorial: When the leadership of Science magazine does not seem to care about getting the facts right in a news story...

A very troubling incident occurred some years ago and I think the concerns and lessons are important.

My apologies for the length, but it is a complicated story.

a cartoon says hey everybody an old man 's talking while bart simpson looks on

ALT: a cartoon says hey everybody an old man 's talking while bart simpson looks on

media.tenor.com

In the summer of 2021, I was contacted by a communications staff member at Pitt with a question about the practices at Science magazine (where I had been Editor-in-Chief from 2016 to 2019).

2/n

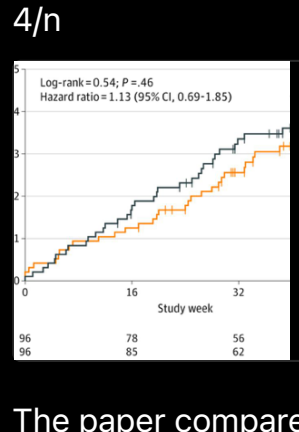
The question related to an interaction between a news reporter from Science magazine (Charles Piller) and a pediatrics faculty member at Pitt (Dr. Juan Celedón).

3/n

Piller was working on a story and had sent Dr. Celedón a list of 28 questions. These were related to a paper published in JAMA in 2020 about a clinical trial of using high doses of vitamin D as a treatment for asthma in children who had low levels of vitamin D.

jamanetwork.com/journals/jam...

4/n



Effect of Vitamin D3 on Severe Asthma Exacerbations in Children With Asthma and Low Vitamin D Levels

This randomized trial compares the effects of vitamin D3 vs placebo on time to severe exacerbation in children with asthma and low vitamin D levels.

jamanetwork.com

The paper compared children with asthma randomized to receive high doses (4000 IU/day) with those who received a placebo. The families of all children were provided a list with foods rich in vitamin D.

5/n

The trial did not show benefits of the high doses of vitamin D on the frequency of severe asthma exacerbations. Indeed, the trial was discontinued early by the data safety monitoring board (DMSB) when it became evident that no statistically significant benefit of high doses would be observed.

6/n

Piller is an investigative reporter at Science who tends to write stories about medical research, usually covering multiple examples related to a topic in his stories. I told the staff member that Piller would likely be covering multiple examples and not focusing just on the Celedón paper.

7/n

The questions covered 2 major topics: (1) The use of a placebo (no vitamin D supplement) in the control group and the risk of bone fractures; and (2) Recruitment of subjects including consent forms and topics related to the racial distribution of the subject population.

8/n

Curiously, the questions also singled out a 2017 paper that had studied vitamin D supplements in children that had not been cited in the JAMA paper.

9/n

I spoke briefly with Dr. Celedón (whom I did not know) and shared my perspectives. He told me that the author of the 2017 paper had been very aggressive in criticizing the JAMA paper in many different fora for years, including repeatedly calling the institutions involved with the study.

10/n

I emailed Tim Appenzeller, the news editor at Science, and let him know that it appeared that a major source for their upcoming story seemed to have a substantial agenda. Appenzeller thanked me for the heads-up.

11/n

A few months later, the article appeared in Science entitled "Failure to Protect: A study of asthmatic children, most of them Black, shows how a common clinical trial design can expose vulnerable participants to serious risks"

www.science.org/content/arti...

12/n

Critics say a childhood asthma study unethically withheld care—and see a troubling trend

Science investigation reveals concerning details of Vitamin D supplement trial that a whistleblower scientist calls "stunning and callous"

www.science.org

I skimmed the article and then started reading it carefully. A few paragraphs in, I read:

13/n

At least 9 kids, across both arms of the trial, broke bones during the trial—nearly double the number expected among asthmatic children over a comparable period. The fractures were (not) disclosed as possible adverse events when the study was published in JAMA last year...

14/n

This sounded quite serious. But it also triggered something in my memory. Weren't fractures mentioned in the questions from Piller that the Pitt communications staff person had shared with me?

15/n

I pulled up the document and looked. Question 9 included a quote from Dr. James Kiley, an NIH official (Director of the lung division at NHLBI) who had helped oversee the trial. This quote included the following:

16/n

"There has been a total of 9 bone fractures occurring after children were assigned to a treatment group – 5 in treatment arm and 4 in the placebo arm. In its analysis... the DSMB did not identify any safety concerns regarding the nature or number of fractures overall or by treatment group..."

17/n

I became quite angry. It appeared that Piller had written about the fractures but had failed to note that the same number of fractures occurred in the treatment and placebo groups. This obviously undermined the argument that the no vitamin D placebo had been responsible for the fractures.

18/n

homer simpson from the simpsons is making a funny face with his mouth open and his fist in the air .

ALT: homer simpson from the simpsons is making a funny face with his mouth open and his fist in the air .

media.tenor.com

I emailed Appenzeller, Holden Thorp, my successor as Editor-in-Chief at Science, and the CEO of AAAS, saying that this was, in my view, unethical behavior by the reporter since he knew about the fracture distribution but did not note this in the article.

19/n

Then I went back to reading. 26 paragraphs after the first mention of the fractures, I saw: "Five fractures had occurred among kids given vitamin D and four in the placebo group, which is nearly twice the rate expected for asthma sufferers in that age group."

20/n

I felt a bit sheepish for my hasty email and I wrote back to Appenzeller and apologized. But I was still quite troubled.

This was still terrible journalism, putting something ominous out there and then not providing key exculpatory evidence until much later in the article.

21/n

However, at this point, my skepticism about this article was quite high. I decided to try to check some of the key statements.

22/n

The first thing I looked into was the rate of fractures. Nine fractures among 192 kids (average age 10 years) over the course of a year did not seem high to me, thinking back to kids in casts in my elementary school days.

23/n

I dug into the literature and found a number of studies in different populations. Based on these studies, the best estimate for the fracture risk is 4% per year. Thus, one would expect 8 fractures among 192 kids, completely in line with the 9 observed fractures.

24/n

Thus, both the number of fractures and the distribution of fractures between the treatment and control groups are consistent with these being normal incidental fractures in 10-year-old children. This was apparently also the opinion of the DMSB which had access to more detailed information.

25/n

According to Dr. Kiley (from Piller) "The DSMB assesses whether there is a statistically different pattern of events between the intervention and control arms of the study. Broken bones would be reportable as adverse events to these oversight bodies. No safety concerns were identified."

26/n

Note that fractures were not defined as "severe (serious) adverse events" in this trial. This is a term with a specific meaning. Because fractures were not defined as severe adverse events, one would not expect them to be reported in publications and other reports.

27/n

The story's key point was the use of a placebo, as opposed to a low vitamin D dose, for the control group.

28/n

Piller wrote: The placebo was justified because vitamin D testing is not routine, they argued. If not for the trial, the kids' vitamin deficiency probably wouldn't have been detected, so they were not worse off in the study."

29/n

However, if one reads the research protocol (available as supplementary info to the JAMA paper), there is a craftily disguised section entitled "Rationale for Placebo in the Control Arm". This does not begin with the Piller rationale. Rather it describes a pilot study involving 24 children.

30/n

Some of these children who were given 200 IU vitamin D/day (vs 4000 IU vitamin D/day for the treatment group in the trial) were found to have serum vitamin D levels above 30 ng/ml. These levels were sufficiently high that they might overlap with the treatment group...

31/n

The discussion provides more details about the expected impacts on children with different vitamin D levels. A discussion of public recommendations follows. Thus, rather than the seemingly cavalier rationale suggested by Piller, the study protocol provides a scientific and scholarly one.

32/n

This does include a discussion of the fact that vitamin D testing in children is not routine.

~9% of the US pediatric population are vitamin D deficient (~7.6 million children). If untreated vitamin D deficiency represented a significant risk, extensive testing would be done.

33/n

The second issue was the racial makeup of the study population. The story repeated noted that more than half of the participants (100/192) of the participants were Black and made numerous statements that implied that this was due to exploitation of underprivileged and vulnerable populations.

34/n

In order to be eligible for the study, kids had to have asthma and to have low vitamin D levels. Dr. Kiley noted that both asthma and low vitamin D levels have a higher prevalence in Blacks than whites. This statement is supported by the literature.

35/n

According to the CDC, the rate of asthma in Black children have asthma at a rate that is 1.9 times higher than that for Black children. This is a substantial factor that could lead to great representation of the Black children in the study.

36/n

Most of the active form vitamin D in the body is produced by sunlight exposure to the skin. Skin pigments interfere with this process. Thus, low vitamin D levels are more prevalent in Black children than in white children by a factor of approximately 3.

37/n

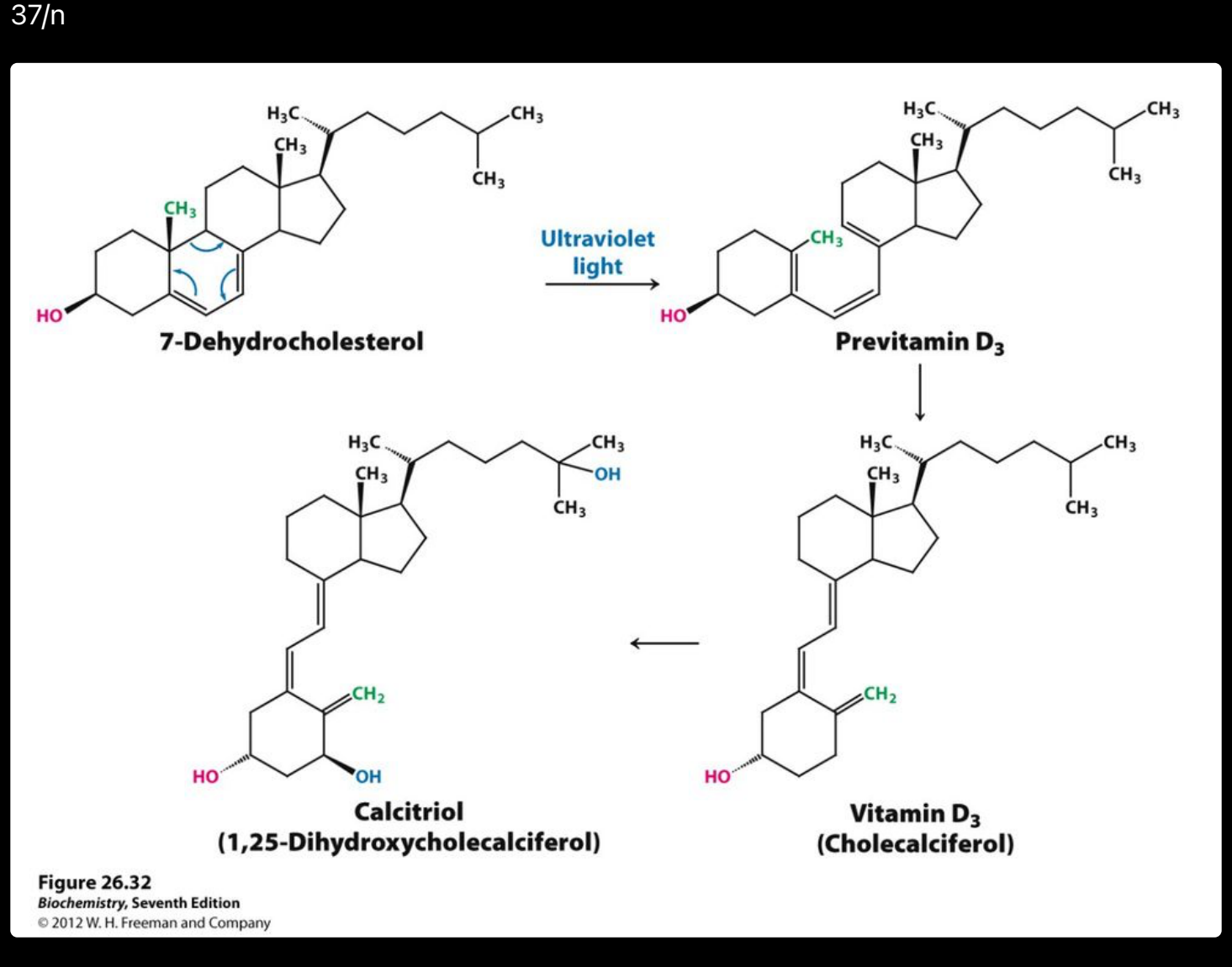


Figure 26.32
Biochemistry, Seventh Edition
© 2012 W. H. Freeman and Company

Thus, the combination of risks for the two inclusion criteria is approximately 2 X 3 = 6 times higher for Black children compared to white (assuming the risk factors are independent).

38/n

The study was a multicenter trial in 7 cities: Pittsburgh, St. Louis, Cleveland, Boston, Cincinnati, Denver, and San Francisco. Using census data and assuming that each site contributed the same fraction of participants, the percentage of Black eligible participants is anticipated to be 57%.

39/n

The assumption that each site contributes the same number of participants is certainly not correct. Simulations assuming variable levels leads to a 95% confidence interval of 46-68%. Thus, the level of Black participants was consistent with the eligible population at the trial sites.

40/n

Piller's story quotes ethicist Harriet Washington "The hunt for genetic explanations above social linkages "reinforces the belief that ... biological dimorphism drives a lot of illnesses," she says. "It's unethical. It's deeply illogical. And it fits a racist pattern."

41/n

Social factors are absolutely central to health disparities in the United States. But the biological factor of skin color directly affects the topic under study here, namely vitamin D deficiency. Piller pursues an accusation of racism, perhaps because of his lack of biochemical knowledge.

42/n

Pilleer quotes Raymond Givens who studies institutional racism.

"[If] most of subjects will be nonwhite, and a large proportion low-income and perhaps lacking advanced education among parents, there is a need for heightened attention to the ethics and appropriateness of the trial"

43/n

Did the subjects of the study have parents who lacked advanced education? Generally, they did not. The educational attainment of the parents is tabulated in the JAMA paper: 19% had postgrad education, 37% completed college, 23% had some college, and 22% had high school or less.

44/n

To summarize the findings based on literature searches and reading the published study protocol, I found that:

45/n

(1) The number of fractures matched that expected with no evidence of any increase due to the placebo control.
(2) The use of a placebo was thoughtfully examined and justified by the need to ensure that there were differences in vitamin D levels between the treatment and control groups.

46/n

(3) The fraction of Black children enrolled in the study matched the fraction expected from the prevalences of asthma and low vitamin D levels and the demographic characteristics of the study sites.

47/n

(4) There was publicly available information that the participants as a rule did not come from family with low educational attainment

48/n

These findings undermine most of the themes of the story. In addition to my general concerns about the needs for accuracy in the media, my distress was accentuated by the fact that the article covered important and timely topics.

49/n

First was the point of whether and when using a placebo control was appropriate. But, in addition, the story touched on the participation of both children and Blacks in clinical trials and conveyed both explicit and implicit messages that the trial participants had been abused and exploited.

50/n

I was quite distressed by these findings. However, things got substantially worse when I found out that Science had done a podcast with Piller and I listened to it.

51/n

The podcast involved an interview between the podcast host (Sarah Crespi) and Piller. Early on, the interviewer asked about the fractures in the study. Here is Piller's response

(SOUND ON)

He said that the 9 fractures VASTLY (his emphasis) exceeded the number expected. The article said that the number of fractures was "nearly double the number expected among asthmatic children over a comparable period."

53/n

Even if you accept the statement about the expected number, 9 compared to an expectation of 5 in a population of 192 children is not statistically significant (p = 0.42).

Vastly, hardly.

54/n

He also referred to these as "severe adverse events" which, as I noted previously, is a technical term with a specific meaning defined in the trial. These usually include things that require hospitalization and there was no evidence these the incidental fractures approached this.

55/n

But it then got much worse...

SOUND ON

The interviewer continued: "Because these fractures weren't described in any detail, we don't know if more occurred in the placebo group or the high vitamin D group. We don't know if they're from low impact walking or high impact sports, could something else be happening here?"

57/n

The interviewer had apparently missed the fact that this information was in the story. I couldn't blame her too much for this as I had missed it as well since it was buried 26 paragraphs after the fractures were first mentioned.

58/n

I assumed Piller would point that out. But he said
"The biggest likelihood is that the low-dose may have had a deleterious effect"

The interviewer corrected him "The no-dose."

Piller responded "Yes. The no-dose had a deleterious effect on bones, that's the major risk in this age group."

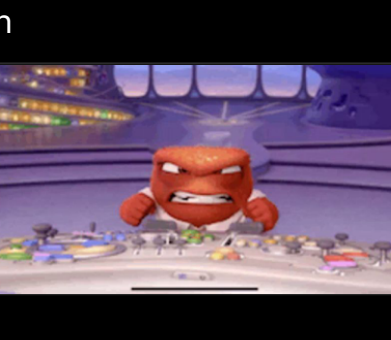
59/n

The Interviewer said: "But we shouldn't be guessing."

Piller said: "We shouldn't be" and did not mention that he knew (and wrote about it).

WTAF

60/n



a cartoon character with an angry face is sitting at a desk

ALT: a cartoon character with an angry face is sitting at a desk

media.tenor.com

I was greatly alarmed by this clear breach of journalistic ethics. I listened to the rest of the podcast and then listened to it again. I sounded even worse to me the second time and there were other uses of inflammatory language that I will not detail here.

61/n

I emailed my successor as Editor-in-Chief at Science, Holden Thorp, Appenzeller, and the CEO of AAAS with my concerns about the story and, particularly, the podcast. In response, they posted a "correction" on the page of the Science podcast website. This read:

62/n

"In the Vitamin D trial segment, the host says it is unknown if bone fractures appeared in placebo or treatment groups. In fact, while this information was not disclosed in a peer-reviewed publication, this is known through details submitted to Congress—please see the full story for details."

63/n

This, in my view, made the situation worse. They did not acknowledge that more fractures occurred in the treatment than the control group. Instead, they doubled down on the fact that, in their view, the occurrence of these fractures had been inappropriately hidden from the public.

64/n

I will not detail the many additional steps that took place, but the Science leadership never acknowledged the errors in the story. They did publish Letters-to-the-Editor from me and Dr. Celedón and even gave us more than the normal amount of space.

www.science.org/doi/10.1126/...

65/n

Editor's note

The article by Charles Piller raises important questions about clinical trials that compare experimental treatments against a placebo instead of standard care by physicians—a topic of active debate in...

www.science.org

The response (from Appelzeller) makes 3 points:

"...multiple ethicists and experts in clinical trial design who reviewed key trial documents expressed the same concerns (as the primary source)".

Were these experts informed about nuanced rationale for placebo control including the pilot study?

66/n

"Piller asked for details on the bone fractures and why they were not publicly reported as adverse events. Celedón elected to respond to Piller's questions with a short general statement."

67/n

I am sorry if my advice regarding the process at Science led Dr. Celedón not to respond more fully.

But, Appenzeller does take into account that Dr. Celedón had been dealing with harassment from Piller's source for years.

68/n

Appenzeller goes on
"On the day Science published the article, the fractures appeared on ClinicalTrials.gov as adverse events."

69/n

The innuendo that updating ClinicalTrials.gov when information about the trial had been made public by others reflects some sort of admission of negligence in the past is disgraceful in my opinion.

70/n

Appenzeller closes "they avoid discussing an ethical concern shared by experts quoted in the article, which is that any risks of the trial fell predominantly on Black children."

71/n

The minimal risks on the trial fell on the group would had the highest stake in the trial and would benefit the most from any findings. This is appropriate and did not reflect racist behavior despite the unfounded accusations in the story.

72/n

I was greatly disappointed and distressed that the leadership at Science never acknowledged their factual errors regarding the occurrence of fractions and about the racial distribution of the study population.

73/n

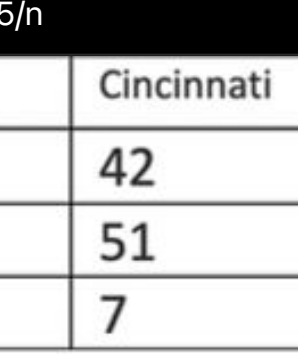
It is important that journalists, particularly science journalists, and especially journalists at places like Science and Nature are careful about using quantitative information carefully and precisely. However, in this case, the leadership at Science does not appear to care about this.

74/n

I wrote an essay published on Medium with details of many of the arguments here to emphasize this point.

@jeremymberg/the-importance-of-careful-use-of-comparative-terms-in-quantitative-arguments-in-science-journalism-4c95e75c7729" target="_blank">medium.com/@jeremymberg...

75/n



Cincinnati
42
51
7

The importance of careful use of comparative terms in quantitative arguments in science journalism

The credibility of all journalism rests on careful and responsible use of verifiable results in presenting the stories told. The...

medium.com

I wrote to Holden Thorp some time ago to try to move on from this sordid affair. His response:

Tim, Charlie, and I have been over all of this multiple times and are comfortable with where we came out.

76/n

All-in-all, the situation reminds me a bit too much of the Washington Post and. Jeff Bezos. I have a great deal of respect and admiration for most of the people at Science.

However, when the leadership does not have clear enough ethical sense to admit their own mistakes, it is deeply problematic.

In some senses this situation is worse as Bezos was trying to curry favor from the incoming President of the United States while the leadership at Science seems to be merely "circling the wagons" to protect a reporter who wrote as deeply flawed story, or for some other less obvious reason.

/fin

a close up of a man wearing a shirt and tie

ALT: a close up of a man wearing a shirt and tie

media.tenor.com

Some additional comments:

The major source for this story had published a paper in 2017 on treatment of asthma with high doses of vitamin D. However, this study was quite different both in terms of treatment and patient population.

1/n

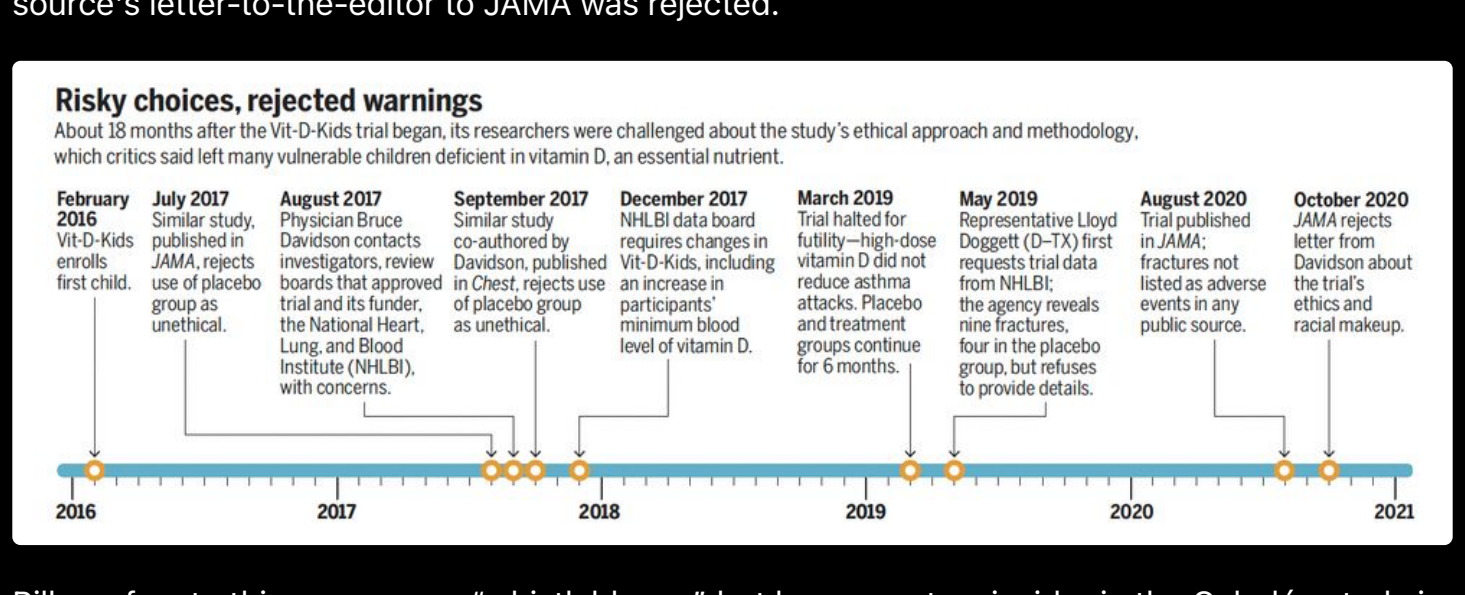
The treatment was one-time injection of 300000 to 600000 IU of vitamin D and then maintenance with oral vitamin D compared with a low dose oral vitamin D control.

2/n

The study was performed in Qatar on children the ER. The median age of the children was 6.3. They were sicker than the patients in the Celedón study with 91% receiving a course of oral steroids every MONTHcompared with a median of 1 such course per YEAR for children in the Celedón study.

3/n

This is the paper that Piller asked about, singling in out among many papers on vitamin D and asthma. His story includes a timeline with the publication of this paper and notes that the source's letter-to-the-editor to JAMA was rejected.



Piller refers to this source as a "whistleblower" but he was not an insider in the Celedón study in any way. Perhaps for heartfelt concerns about ethics or, perhaps, for other reasons, he took great issue with the Celedón study when he found out about it.

4/n

He went to the press and to Congress to raise concerns. He was the one who obtained the materials from Congress about the fractures, the DSMB, and so on. He was a valuable source because he could provide this information to Piller.

5/n

Pillar held him out as an expert on vitamin D and asthma. As of 2 years ago, Pubmed searches revealed that he had published 1 paper (the 2017 paper) and 3 commentaries on vitamin D over a long career. He published 6 papers of asthma. He had published results on clinical trials on other topics.

6/n

In contrast, Dr. Celedón had published 62 papers on vitamin D, 255 papers on asthma, and 20 papers of health disparities or health equity.

7/n

The source expressed strong and inflammatory opinions about the Celedón trial. While the Piller story was in process, the source posted a comment on Twitter (in response to a story about anti-racism):

8/n

BLDMDMPH @bldmdmph · Jul 11

And when the new US government-sponsored Tuskegee study is exposed next month, will Mr Staples and the NYT cover it?

He also posted when the story came out (with some inflammatory comments):

BLDMDMPH @bldmdmph · Aug 21

"No one cares about the kids." That's modern American medical research too, particularly minority kids, as reported 13 August in Science, the 21st century Tuskegee study:

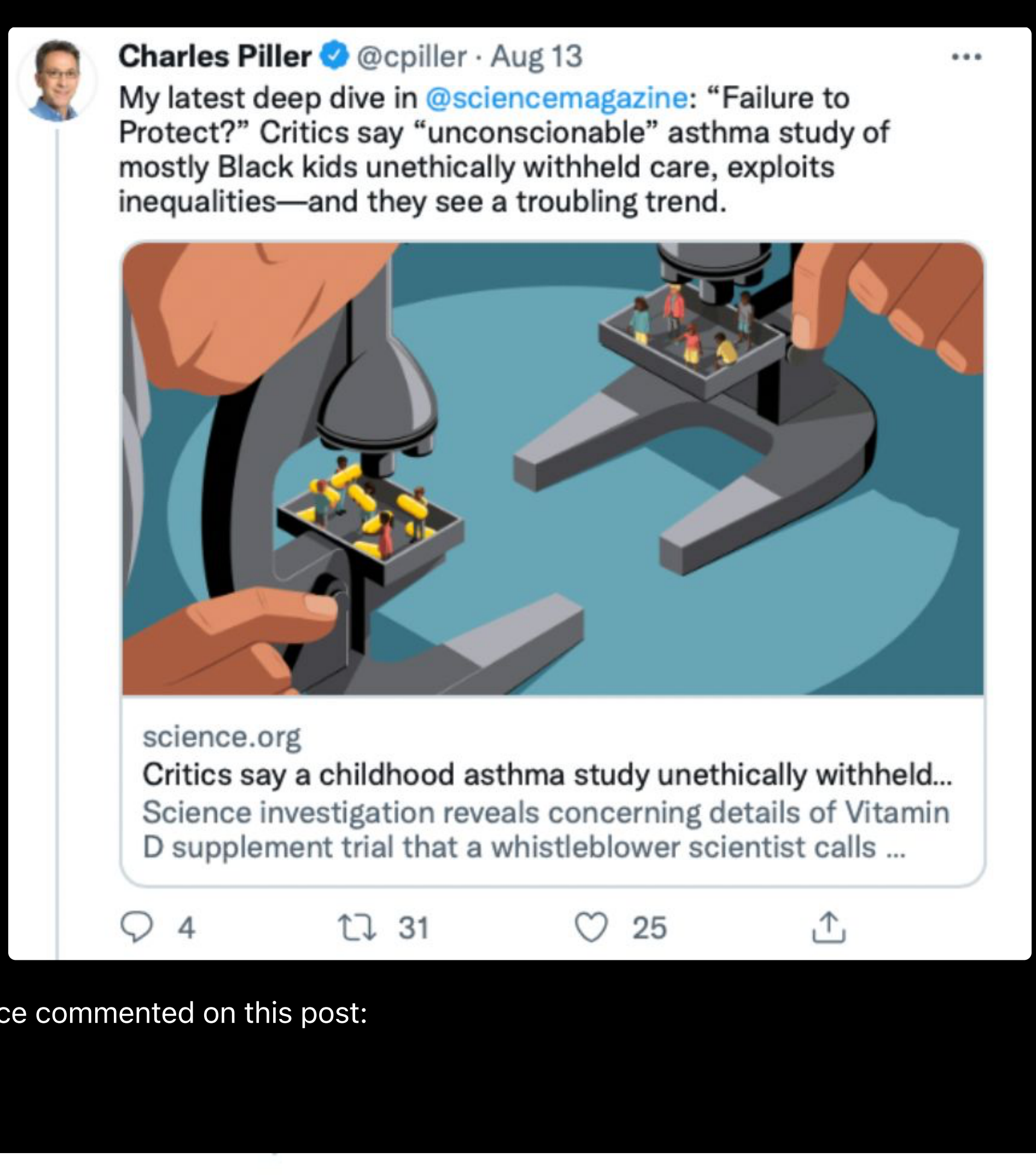
science.org

Critics say a childhood asthma study unethically withheld...

Science investigation reveals concerning details of Vitamin D supplement trial that a whistleblower scientist calls ...

Piller also posted his story:

10/n



Charles Piller @cpiller · Aug 13

My latest deep dive in @sciencemagazine: “Failure to Protect?” Critics say “unconscionable” asthma study of mostly Black kids unethically withheld care, exploits inequalities—and they see a troubling trend.

science.org

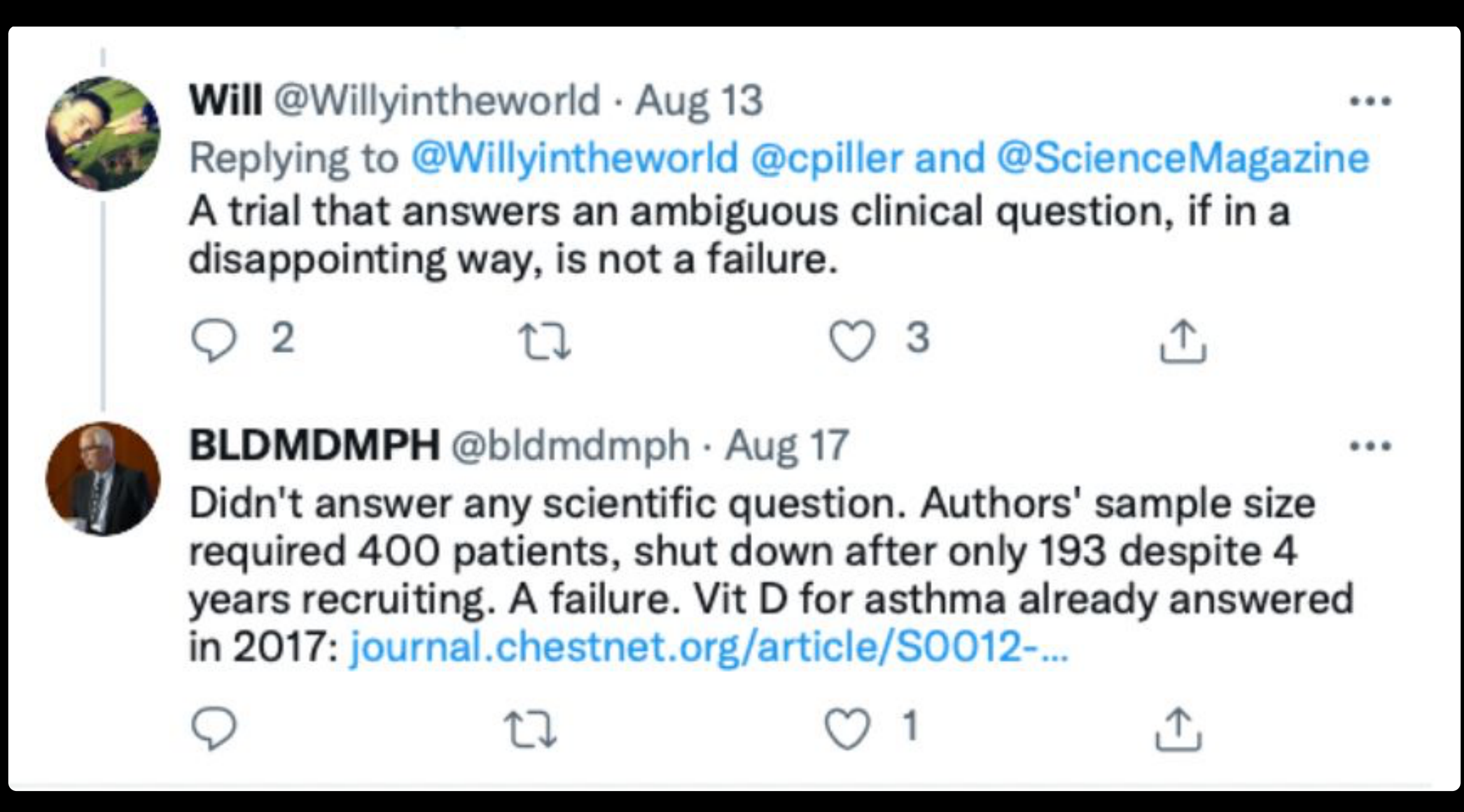
Critics say a childhood asthma study unethically withheld...

Science investigation reveals concerning details of Vitamin D supplement trial that a whistleblower scientist calls ...

4 31 25

The source commented on this post:

11/n



Will @Willyintheworld · Aug 13

Replying to @Willyintheworld @cpiller and @ScienceMagazine

A trial that answers an ambiguous clinical question, if in a disappointing way, is not a failure.

2 3

BLDMDMPH @bldmdmph · Aug 17

Didn't answer any scientific question. Authors' sample size required 400 patients, shut down after only 193 despite 4 years recruiting. A failure. Vit D for asthma already answered in 2017: journal.chestnet.org/article/S0012-...

1

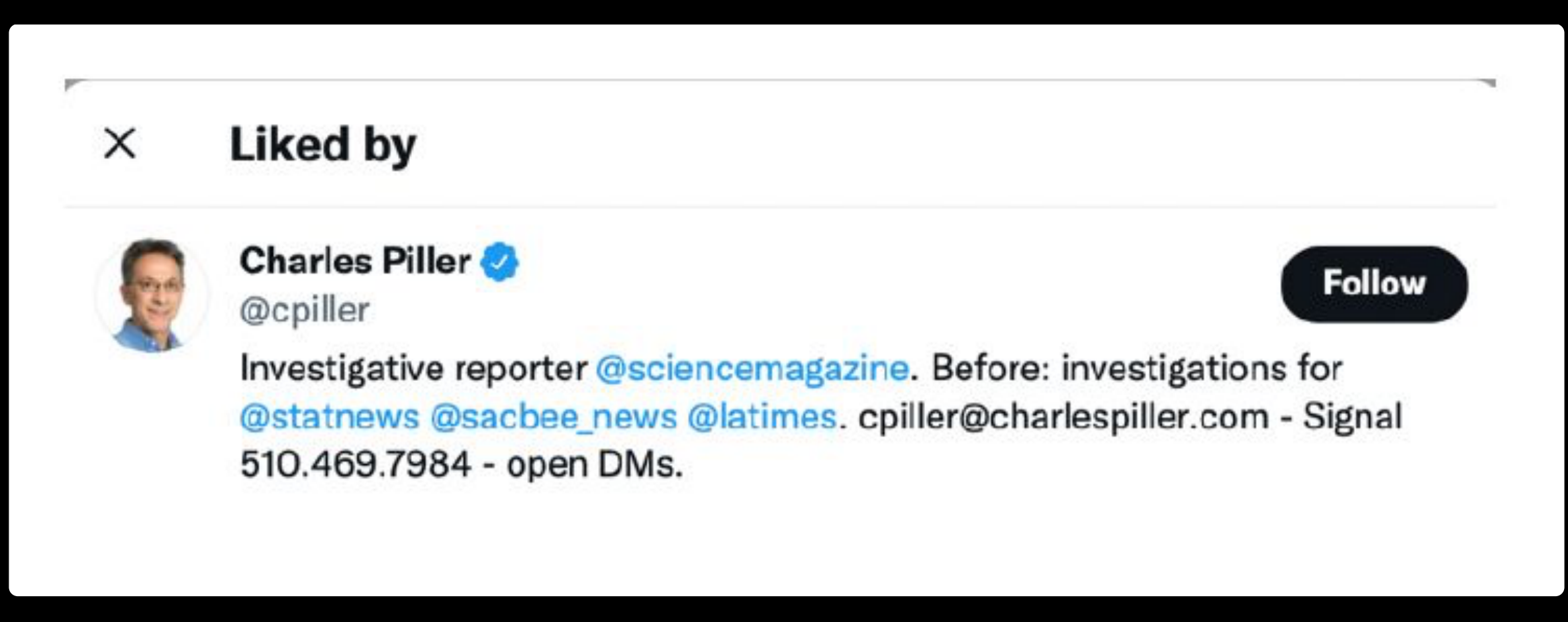
This comment asserts that his 2017 paper settled the issue. This is absurd since the study was quite different and since a single trial rarely settles a complicated clinical issue.

Moreover, as the comment above notes, a trial that provides clear evidence, even if negative, is not a failure.

12/b

The comment received 1 "like" (from Piller)

13/n



Liked by

Charles Piller

@cpiller

Investigative reporter @sciencemagazine. Before: investigations for @statnews @sacbee_news @latimes. cpiller@charlespiller.com - Signal 510.469.7984 - open DMs.

This behavior suggests a fairly cozy relationship between Piller and his source.

When I posted my article on Medium, I shared the link the people at Science. I was surprised that my article received a comment quite quickly. This was even before the article was discoverable on Google.

14/n

Early in the comment, it reads:

A colleague brought my attention to this "quantitative argument" rebuttal by this PhD Pittsburgh colleague of the criticized study principal investigators, who himself has no clinical trial recruitment experience nor even a medical degree.

15/n

Either the source has a "colleague" who is an avid Medium reader or this "colleague" was Piller.

The source concludes his comment about my article with:

(This article) recalls the wise aphorism: "Figures don't lie but.." You know the rest of it.

16/n

The Piller story includes some other material:

He quotes physician Michael Carome, a former top regulator at OHRP who directs health research for the consumer advocacy group Public Citizen.

17/n

His is quoted:

The trial protocol noted the high-dose supplement was tested against a placebo to avoid a “false negative” outcome. “They wanted to maximize their chances of finding a difference."

18/n

This seems to me to be an important reason for the choice as doing a trial that had little chance of finding a difference is clearly unethical.

Piller continued to quote Carome:

19/n

To Carome, Vit-D-Kids offers new evidence that, overall, “our IRB system is broken.” He doubts that any hospital panel that greenlit the study asked how its own physicians would normally treat asthmatic children deficient in vitamin D. None would fail to give baseline supplements, he says.

20/n

He goes on...

"It’s a sign that IRBs tend to uncritically accept NIH funding as a stamp of approval"

To say that this comment is inconsistent with my experience is an understatement.

21/n

Based on observations of IRBs at multiple institutions and discussions with hundreds of clinical investigators, IRBs certainly do not accept NIH funding (or much else) uncritically.

IRBs work hard zealously to protect patients and to make sure that proper processes are followed.

/fin

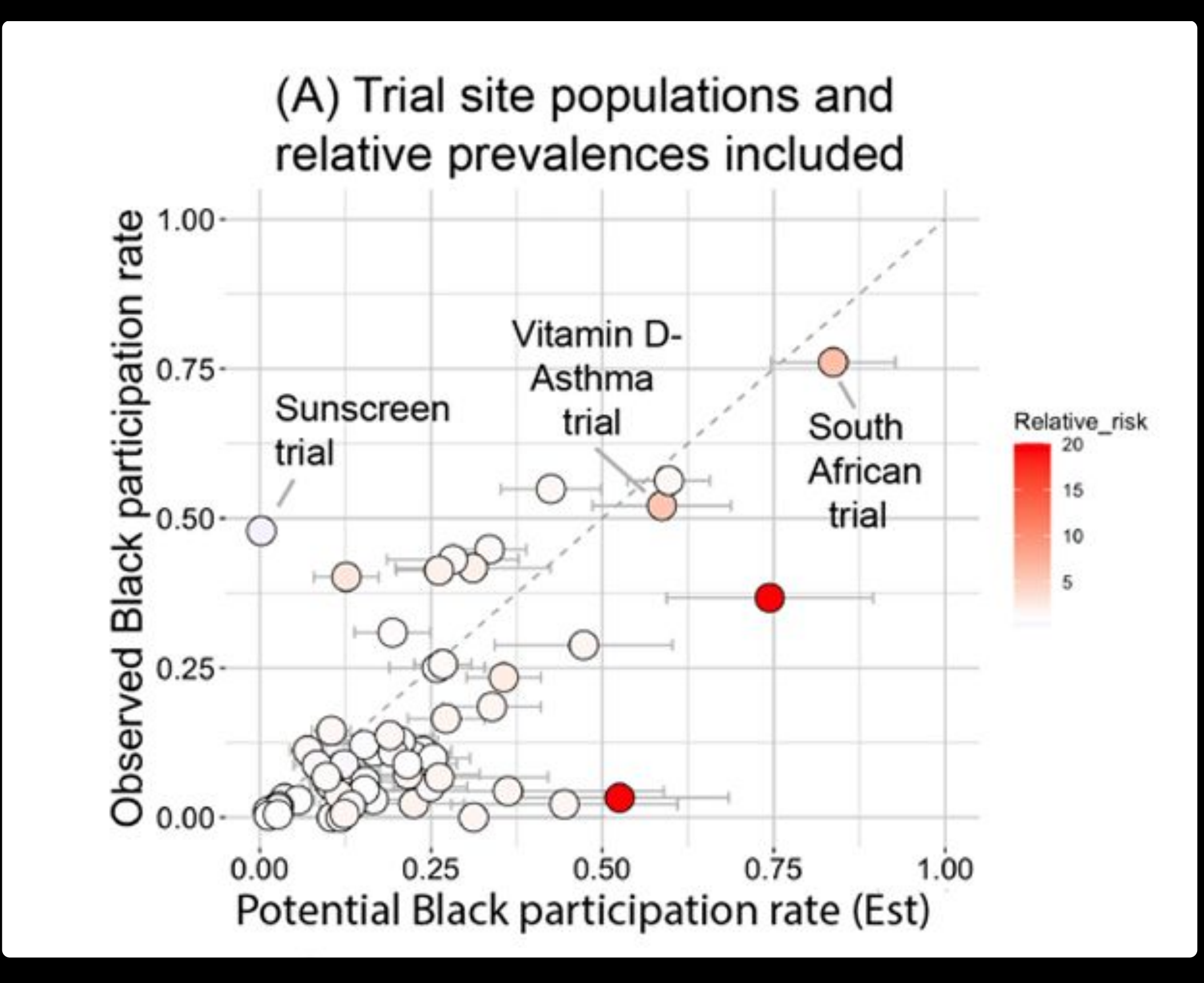
I got interested in whether Blacks were represented at appropriate levels (based on the condition(s) under study and the demographics at the study site). I looked at all clinical trials published in JAMA, New England Journal of Medicine, and The Lancet in 2020.

1/n

58 out of 359 published trials had sufficient information to allow the sort of analysis that I had used for the Celedón trial.

The fraction of Black participants is compared with that anticipated is shown here.

2/n



The level of Black participation was less than that based on prevalence and demographics in 47 out of 58 trials and the difference of mismatch is a factor of 2 or more for 23 trials.

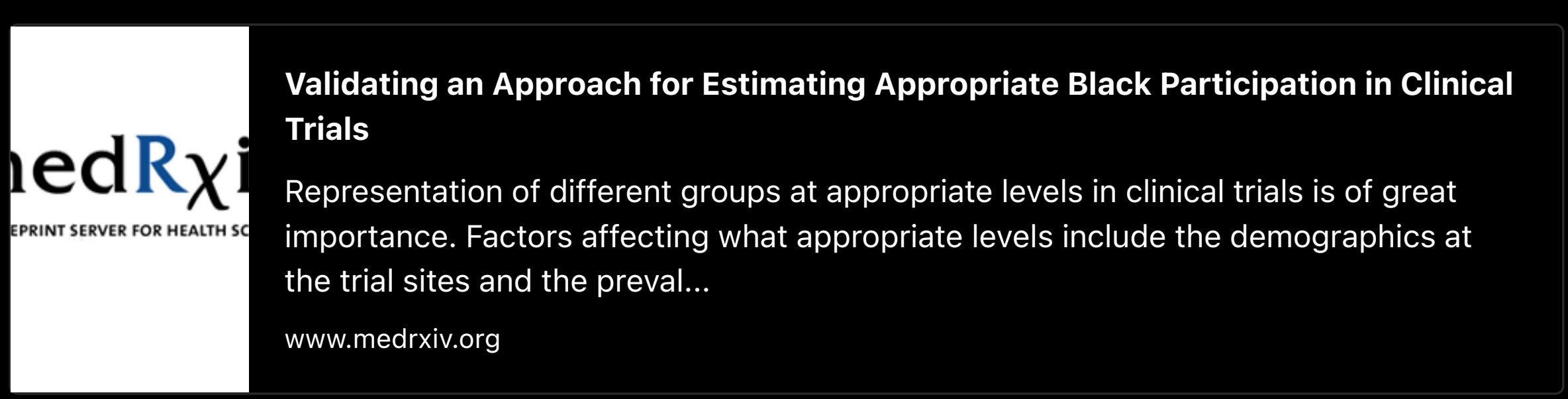
3/n

The values for the vitamin D (Celedón) trial match well as do some large trials such as one of drug-resistant tuberculosis conducted in South Africa.

I wrote an article, co-authored with Dr. Celedón and Dr. Naudia Jonassaint that is posted on MedRXiv.

www.medrxiv.org/content/10.1...

/fin



Validating an Approach for Estimating Appropriate Black Participation in Clinical Trials

Representation of different groups at appropriate levels in clinical trials is of great importance. Factors affecting what appropriate levels include the demographics at the trial sites and the preval...

www.medrxiv.org