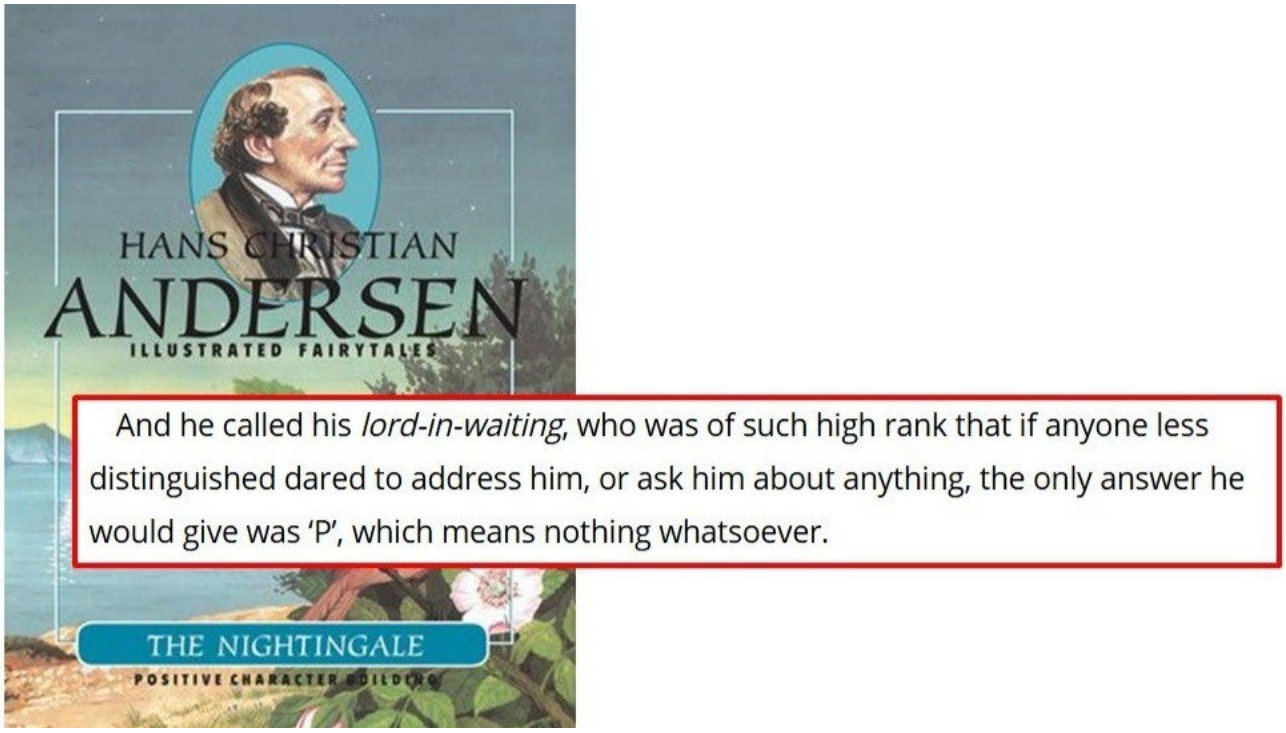

Hans Christian Andersen in "The Nightingale" correctly observed that P is overvalued

Paradigm changing research: Patterns or p-values?

March 2, 2025

By Peter Aaby and Christine Stabell Benn

From time to time, some people attack our research into the non-specific effects of vaccines with arguments that are centered around the p not being <0.05 or lack of methodological rigor. Few of these people have ever studied vaccines. They have apparently learned about the Popperian falsification: one must propose a 0-hypothesis of no difference between groups (in the case of vaccines, two groups that received different vaccines) and reject it with a statistical test that shows a p-value of <0.05 to demonstrate that two groups differ from each other. They have apparently also learned that one cannot believe any study result if there was no hypothesis *a priori* or if it is not the result of the primary analysis.

It is worthwhile to study where such ideas come from and why they persist, though experts warn against such a dichotomized view of p-values (1) and though they are counter-productive to discovery. We will here illustrate how these ideas can hamper progress in paradigm-changing research, using our research as a case.

Changing a paradigm

We will briefly present our discovery of the non-specific effects of vaccines, i.e. the idea that a vaccine can have effects which cannot be explained by the prevention of the infection targeted by the vaccine (2,3).

The discovery was based on a series of observations, which contradicted the current paradigm: that vaccines have only specific effects.

When an old paradigm is no longer tenable, there is a need to develop a new paradigm. We are now in the iterative research process of exploring the new “language” of non-specific effects (Figure 1): when are the non-specific effects present, do they differ in character between different vaccines, and how long do they last? This is a part of the discovery where one is bound to commit errors, because obviously one does not know the new language from the beginning.

Numerous studies have led to the tentative formulation of new generalizations and based on these we make deductions and generate new hypotheses, which we then test (Figure 1).

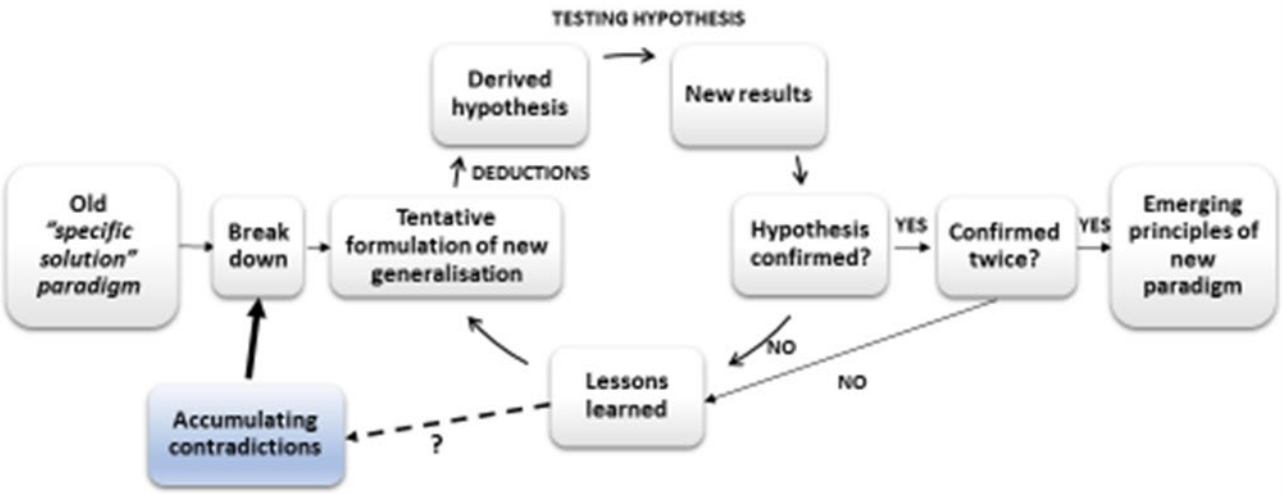

Figure 1. Our iterative research process.

Sometimes we confirm our hypotheses. Provided we have confirmed an observation twice or more, we start to incorporate it in our emerging principles of the new paradigm.

As expected in such a process: S***ometimes we are not able to confirm our initial hypotheses.*** We then try to understand why. Hence, a refutation of one of our hypotheses will usually lead to modification of the generalizations, to new hypotheses, which we then test (Figure 1).

Using measles vaccine as an example of paradigm-changing research

We had seen that measles vaccine had much stronger effects on overall child mortality than could be explained by prevention of measles infection (4-6). Based on this and other observations we hypothesized that it would be beneficial to give measles vaccine early (5). We were able to confirm that it was overall beneficial (6), thus adding to the observations, which contradict the old paradigm of vaccines having only specific effects and point towards new emerging principles. However, we also established that some immune stimulants, like administration of non-live vaccines with or after the measles vaccine, will reduce or remove the beneficial non-specific effects of measles vaccine (7). We also detected that several other live

vaccines appear to have similar beneficial non-specific effects (2,8). The consistency in beneficial non-specific effects, the consistency that non-live vaccines with or after live vaccines will reduce the beneficial effects, and the observation that the pattern of beneficial non-specific effects can be extended to other live vaccines, make it very unlikely that the non-specific effects of measles vaccine should not be true (2).

Using DTP-vaccine as an example of paradigm-changing research

The three doses of diphtheria-tetanus-pertussis (DTP) vaccine have been the backbone of the immunization program in low-income countries. The third dose of DTP (DTP3) has been used as the main performance indicator. It was therefore very surprising, when we started looking at the potential non-specific effects of DTP, that we found that DTP-vaccinated children had higher mortality than DTP-unvaccinated children (Figure 2)(2).

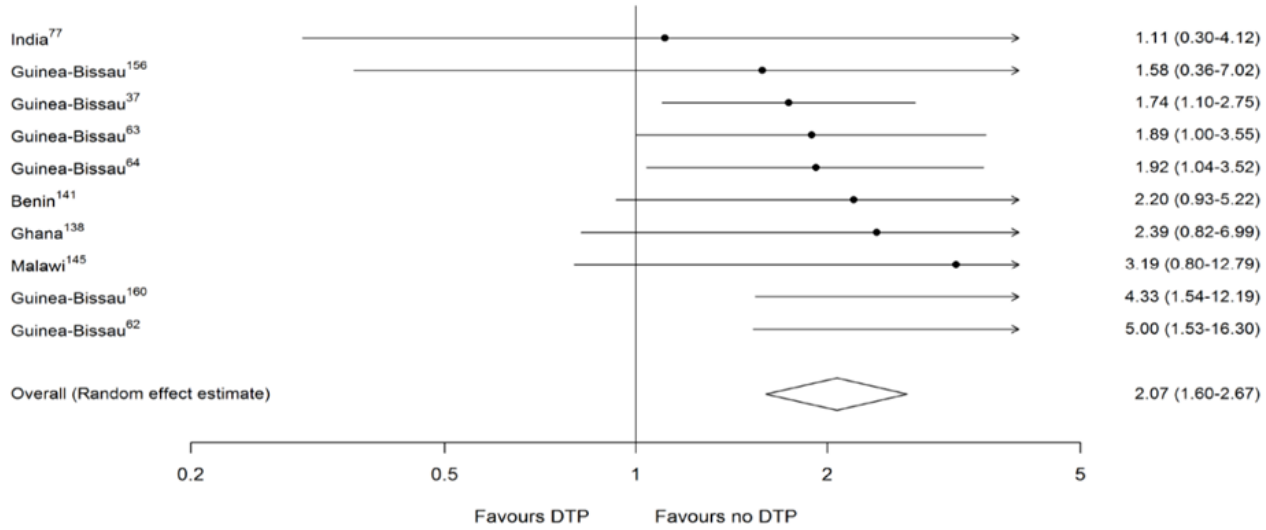

Figure 2. Meta-analysis of the effect of DTP on all-cause mortality in prospective studies with no survival bias (2).

Noteworthy, in Figure 2 some studies were not significant in their own right, because the confidence interval crosses 1. However, the overriding picture is quite clear and consistent: all studies uniformly suggest that being DTP-vaccinated is associated with higher mortality than not being DTP vaccinated. The overall analysis shows that DTP vaccine is associated with a 2-fold increase in mortality (9), which is highly significantly different from 1.0, and even more so from the expected beneficial effect of protecting against diphtheria, tetanus, and pertussis. We also found that this deleterious effect of DTP was particularly pronounced for females (9). Hence, though females have usually slightly lower child mortality than males, after DTP the female-male mortality ratio is above 1.0. We furthermore detected that several other non-live vaccines (N=6) appear to have similar deleterious non-specific effects, particularly for females (2). The consistency in these deleterious non-specific effects, the consistency with which they are more pronounced for females, and the observation that the patterns can be extended to other non-live vaccines make it very unlikely that the deleterious non-specific effects should not be true.

Using this iterative research process, we continuously refine our observations. By now we have 11 emerging principles (Figure 3):

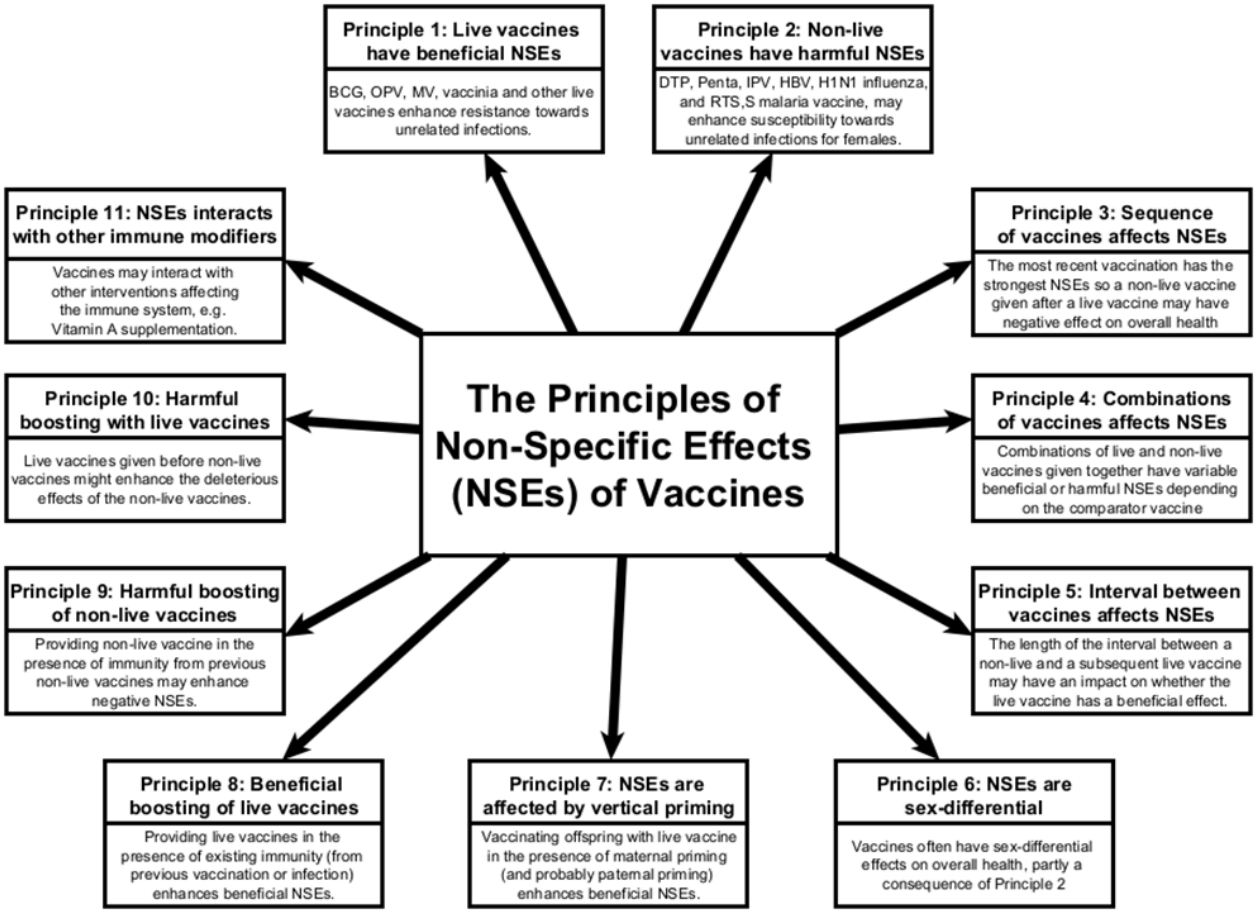

Figure 3. The 11 emerging principles of non-specific effects of vaccines

In this iterative research process, we continuously gather evidence and continuously attempt to identify the overall interpretation of data, which explains all observations. In this universe, p-values in single studies are naturally less interesting than the overall consistency of data.

Discrediting vaccine research

The observations of non-specific vaccine effects have been surprisingly consistent in epidemiological data from low-income countries (2,3). Furthermore, a large amount of immunological data now adds biological credibility by showing that live vaccines may boost innate pro-inflammatory responses that can make the child more resistant against unrelated infections (10), whereas non-live vaccines may induce tolerance that can make the child more susceptible to unrelated infections (11). The discovery of non-specific effects of vaccines is now on Nature's list of major milestones in Vaccines (milestone 13) and the latest version of the

World's leading textbook in vaccines, "Plotkin's Vaccines", has a chapter devoted to non-specific effects, and states firmly that "*it is clear that vaccines have NSE*".

Yet, critique linger. Very few have tried to disprove these patterns with reinterpretations of existing data or with collection of new data. However, many try to discredit such paradigm-changing research for breaking the rules of "proper" research (12-16). The most common approaches are summarized here.

Discredit: The study is not statistically significant, p>0.05.

This is the most commonly heard critique (14-16). For some people, no research result is interesting or worth paying attention to unless the study is statistically significant at a p<0.05 level, studies without such results are considered "null-studies". It is a strange illusion that all studies can maintain statistical power through a prolonged research process and that they should therefore all be statistically significant to be of any importance. This is particularly so for vaccine studies, because once a vaccine has been recommended by WHO it cannot easily be tested directly in a randomized trial, and it has often been necessary to rely on observational studies based on existing datasets.

In the example of DTP, we have used all datasets available where DTP-vaccinated and DTP-unvaccinated children can be compared in a relatively unbiased fashion (2,9). The many studies can then be added in meta-analyses to determine whether the patterns are repeatable.

Hence, the relevant criterion is whether the observations point in the same direction - see for example Figure 2 - not the single p-value in each study.

Discredit: The *a priori* clause: the study was not planned - or the result was not on the "primary outcome".

Critics are often claiming that *post-hoc* analyses are not valid. However, studies generating a new paradigm are unlikely to be *a priori* studies. The more one emphasizes *a priori* the less likely one is to observe anything new and unexpected since it would not be unexpected if it was *a priori*. Sometimes *a priori* is emphasized without realizing the many absurd implications. For example, when we first presented the negative effect of high-titer measles vaccine (HTMV) for female survival in both Guinea-Bissau and Senegal at a WHO meeting in 1991 (17,18), the WHO experts dismissed the observations as not plausible and irrelevant because the studies had been planned for other purposes (19). However, already the following year the observation had been confirmed in Haiti and Sudan and WHO had to withdraw HTMV in 1992 (20).

Studies of child mortality are likely to take 5-6 years or more. Given the current speed of changes in health policies, major changes may take place during a study because new vaccines are introduced, campaigns are implemented, the vaccine sequence or co-administration is changed. If the effect of such events on a study intervention cannot be reported without being called *post-hoc*, we are clearly limiting the possibilities of seeing something new and getting smarter.

Furthermore, some people will dismiss results if they are not related to the primary outcome. E.g., a study showing no effect of BCG vaccine on infant mortality, which was the primary outcome, but a strong protective effect on BCG vaccine on neonatal mortality, a secondary

outcome, has been classified by critiques as a null-study (8). It should be self-evident that this is a recipe for never seeing anything.

Discredit: The study is observational study: It could be bias.

People, who like to maintain status quo in vaccine research and criticize controversial observations, will usually emphasize that the controversial studies have been observational studies and not randomized trials and that they are therefore biased and unreliable (21). Applied to vaccine research this a very dangerous approach.

First, randomized trials vaccines that are already approved by WHO will usually be considered unethical. Demanding randomized trials thus literally means cutting off the possibility to assess these vaccines.

Second, often potential negative effects will have to be studied in observational studies, rather than randomizing children to something which is hypothesized to be dangerous for them.

Importantly, whether an observation is reliable does not depend on whether the study is observational or randomized. Triangulation of observational studies with different underlying confounding structure can be used to build causality arguments through triangulation (22). On the other hand, results from randomized trials are not necessarily correct and reliable. In principle, a trial is only randomized on day 1 (23). Furthermore, there are many examples of other interventions, which have gone unnoticed by the researchers (24), but have affected the outcomes of trials (25).

Discredit: The statistical analysis was not an "intention-to-treat" analysis or did not use a two-sided p-value.

In the statistical analysis of randomized trials, using an intention-to-treat analysis and a 2-sided p-value is natural if one starts from an agnostic basis. However, in the iterative research process, looking for the new language of non-specific effects of vaccines in randomized trials, the most relevant analysis is that of the "per-protocol analysis" where one only considers children, who followed the protocol and got the intended treatment - rather than the "intention-to-treat"(23,26).

Furthermore, it is natural to have a 1-sided p-value, if many prior studies have shown a given effect of a vaccine and the planned study sets out to test that specific effect.

Nonetheless, there are strong proponents for intention-to-treat analysis and 2-sided p-values and they will dismiss results of per-protocol analyses and to dismiss studies with 1-sided p-values.

On normal vs paradigm shifting vaccine research

Most "normal" vaccine research is one-dimensional, studying a vaccine vs no vaccine, placebo or a control vaccine, and examining only the planned outcome. It is implicitly assumed that the vaccine only affects the planned outcome and that no other interventions could affect the association between the vaccine and the outcome. The research is therefore not taking into consideration that other immune stimulants, e.g. other vaccines and vitamins, may modify the

results (24). This is the way the industry tests the vaccines to get regulatory approval. It is very unfortunate that some people perceive this way of doing product development as the only acceptable way of doing vaccine research.

Can we see the unexpected? P<0.05 is not a pattern recognizing receptor

The discrediting strategies usually focus on only one study as being statistically non-significant, *post-hoc*, or observational. There is rarely an attempt to assess the relevance of the critique in relation to all the available studies, the ***totality of data***. The discrediting strategies also tend to ethicize the issues; following the “rules of science” is good - not following them is morally reprehensible: *a priori* is good and *post-hoc* is suspect; $p<0.05$ is the signal of an important association, whereas $p>0.05$ is not significant and there is no association; randomized trials are good and observational studies are bad. However, the essence of science, creating more understanding, is not to have followed the rules but to have found a new pattern that better explains the totality of data than the previous.

The self-perceived scientific rigor has grown among authors, reviewers, editors and critics in the last 40 years where we have examined the non-specific effects of vaccines. The scientific rigor is often justified as being necessary in order to control how the industry is handling data. However, the need to control the industry does not justify that one should not observe what one has not planned to see. All paradigm-changing studies are necessarily unplanned.

As a result, we are getting more and more young researchers, who do not dare to see what they have not planned to see, and who think they are transgressing some sacred rules if they are not following the scheme of protocol, study and reporting of only the planned outcomes, and considering $p<0.05$ a discriminator of important or non-important.

Paradigm shifting research does not come from following rules but rather from following unexpected observations and showing that they can be reproduced in other settings and can generate predictions which are verifiable and enhance the predictability.

References

1. American Statistical Association. Statement of statistical significance and p-values. 2016. https://www.amstat.org/asa/files/pdfs/p-valuestatement.pdf

2. Benn CS, Fisker AB, Rieckmann A, Sørup S, Aaby P. Vaccinology: Time to change paradigm? Lancet Infect Dis 2020;20(10):e274-e283.

3. Aaby P, Benn CS. Developing the concept of beneficial non-specific effect of live vaccines with epidemiological studies. Clin Microbiol Infect. 2019 Dec;25(12):1459-1467.

4. Aaby P, Bukh J, Lisse IM, Smits AJ. Measles vaccination and reduction in child mortality: a community study from Guinea-Bissau. J Infect 1984; 8:13-21

5. Aaby P, Andersen M, Sodemann M, Jakobsen M, Gomes J, Fernandes M. Reduced childhood mortality after standard measles vaccination at 4-8 months compared with 9-11 months of age. Br Med J 1993;307:1308-1311

6. Aaby P, Samb B, Simondon F, Coll Seck AM, Knudsen K, Whittle H. Non-specific beneficial effect of measles immunisation: analysis of mortality studies from developing countries. Br Med J 1995;311:481-485

7. Clipet-Jensen C, Andersen A, Jensen AKG, Aaby P, Zaman K. Out-of-Sequence Vaccinations With Measles Vaccine and Diphtheria-Tetanus-Pertussis Vaccine: A Reanalysis of Demographic Surveillance Data From Rural Bangladesh. Clin Infect Dis. 2021 Apr 26;72(8):1429-1436.

8. Aaby P, Roth A, Ravn H, Napirna BM, Rodrigues A, Lisse IM, Stensballe L, Diness BR, Lausch KR, Lund N, Biering-Sørensen S, Whittle H, Benn CS. Randomized trial of BCG vaccination at birth to low-birth-weight children: beneficial nonspecific effects in the neonatal period? JID 2011;204:245-52

9. Aaby P, Ravn H, Benn CS. The WHO Review of the Possible Nonspecific Effects of Diphtheria-Tetanus-Pertussis Vaccine. Pediatr Infect Dis J. 2016 Nov;35(11):1247-1257

10. Arts RJW, Moorlag SJCFM, Novakovic B, Li Y, Wang SY, Oosting M, Kumar V, Xavier RJ, Wijmenga C, Joosten LAB, Reusken CBEM, Benn CS, Aaby P, Koopmans MP, Stunnenberg HG, van Crevel R, Netea MG. BCG Vaccination Protects against Experimental Viral Infection in Humans through the Induction of Cytokines Associated with Trained Immunity. Cell Host Microbe 2018;23(1):89-100.e5.

11. Blok BA, de Bree LCJ, Diavatopoulos DA, Langereis JD, Joosten LAB, Aaby P, van Crevel R, Benn CS, Netea MG. Interacting, Nonspecific, Immunological Effects of Bacille Calmette-Guérin and Tetanus-diphtheria-pertussis Inactivated Polio Vaccinations: An Explorative, Randomized Trial. Clin Infect Dis 2020;70(3):455-463

12. Fine PEM, Smith PG, Evans SJW. Non-specific effects of BCG? JID 2012;205:515

13. Aaby P, Roth A, Biering-Sørensen S, Ravn H, Rodrigues A, Whittle H, Benn CS. No evidence of bias in trial showing BCG reduces neonatal mortality. JID 2012;205:516-7

14. Strøm C. Dødelighed ved difteri-tetanus-pertussis. vaccine. Ugeskr Laeger. 2024 Dec 2;186(49):V07240483

15. Stensballe LG. Does BCG have non-specific effects on childhood mortality? BMJ. 2017 Feb 10;356:j700.

16. Besançon L, Bik E, Heathers J, Meyerowitz-Katz G. Correction of scientific literature: Too little, too late! PLoS Biol. 2022 Mar 3;20(3):e3001572.

17. Aaby P, Knudsen K, Whittle H, Lisse IM, Thårup J, Poulsen A, Sodemann M, Jakobsen M, Brink L, Gansted U, Permin A, Jensen TG, Andersen H, da Silva MC. Long-term survival after Edmonston-Zagreb measles vaccination in Guinea-Bissau: Increased female mortality rate. J Pediatr 1993;122:904-8

18. Aaby P, Samb B, Simondon F, Knudsen K, Coll Seck AM, Bennett J, Markowitz L, Rhodes P, Whittle H. Sex-specific differences in mortality after high-titre measles immunization in rural Senegal. Bull WHO 1994;72:761-770

19. Expanded Programme on Immunization. Safety and efficacy of high titre measles vaccine at 6 months of age. Weekly Epidemiol Rec 1991;66:249-251

20. Expanded Programme on Immunization (EPI): Safety of high titre measles vaccines. Wkly Epidemiol Rec 1992;67:357-61.

21. Higgins J, Soares-Weiser K, López-López J, et al. Association of BCG, DTP, and measles containing vaccines with childhood mortality: systematic review. BMJ 2016;355:i5170.

22. Lawlor DA, Tilling K, Davey Smith G. Triangulation in aetiological epidemiology. Int J Epidemiol. 2016 Dec 1;45(6):1866-1886

23. Hernán MA, Hernández-Díaz S. Beyond the intention-to-treat in comparative effectiveness research. Clin Trials. 2012 Feb;9(1):48-55.

24. Nielsen S, Möller S, Benn CS, Aaby P. The importance of quality of health campaign information. A case study from Guinea-Bissau and Bangladesh. Vaccine X 2024;21:100588

25. Aaby P, Nielsen S, Fisker AB, Pedersen LM, Welaga P, Hanifi SMA, Martins CL, Rodrigues A, Chumakov K, Benn CS. Stopping Oral Polio Vaccine (OPV) After Defeating Poliomyelitis in Low- and Middle-Income Countries: Harmful Unintended Consequences? Review of the Nonspecific Effects of OPV. Open Forum Infect Dis. 2022 Jul 27;9(8):ofac340.

26. Tripepi G, Chesnaye NC, Dekker FW, Zoccali C, Jager KJ. Intention to treat and per protocol analysis in clinical trials. Nephrology. 2020; 25: 513–517.