testing-focused research or possible causal mechanisms that could lead to changes in incidence. Here I try to meet this latter need.

I have adduced very substantial quantities of data to support the hypothesis that mammalian (including human) sex ratios (proportions male) at birth are partially controlled by the hormone levels of both parents around the time of conception.²⁻⁵

Baron-Cohen⁶ hypothesized that one cause of autism is high maternal intrauterine concentrations of testosterone. If both hypotheses were correct, then sibs of autistic probands should contain an excess of brothers. I reported evidence for such an inference at a statistically significant level,⁷ and this finding has itself been confirmed.⁸ So there is good evidence for Baron-Cohen’s hypothesis that one cause of autism is high maternal intrauterine testosterone concentration.

There is good reason to suppose that over the past few decades, there has been an increase in women’s testosterone levels. Such an increase would be expected to have been associated with the rise in rates of diabetes,⁹,¹⁰ and of obesity¹¹ in women. Thus, it seems reasonable to propose that the reported increase in rates of autism is at least partially real and due to increasing maternal intrauterine testosterone levels.

References

2 James WH. Evidence that mammalian sex ratios at birth are partially controlled by parental hormone levels at the time of conception. J Theor Biol 1996;¹⁸⁰:271–86.
3 James WH. Further evidence that mammalian sex ratios at birth are partially controlled by parental hormone levels around the time of conception. Hum Reprod 2004;¹⁹⁹:1250–56.
4 James WH. Evidence that mammalian sex ratios at birth are partially controlled by parental hormone levels around the time of conception. J Endocrinol 2008;¹⁹⁸:3–15.
5 James WH. Further evidence for the hypothesis that parental hormone levels around the time of conception are associated with human sex ratios at birth. J Biosoc Sci 2008;⁴⁰:855–861.
7 James WH. Further evidence that some male-based neurodevelopmental disorders are associated with high intrauterine testosterone concentrations. Dev Med Child Neurol 2008;⁵⁰:15–18.

Re: A further plea for adherence to the principles underlying science in general and the epidemiologic enterprise in particular

From MICHAEL HAUPTMANN¹* and CECILE M RONCKERS²

¹Department of Bioinformatics and Statistics, Netherlands Cancer Institute, Amsterdam, The Netherlands and ²Dutch Childhood Oncology Group (DCOG-LATER), The Hague, The Netherlands

*Corresponding author. Nederlands Kanker Instituut, Plesmanlaan 121, 1066 CX Amsterdam, The Netherlands.
E-mail: m.hauptmann@nki.nl

Boffetta et al.⁴ suggest that the participation of one of us (M.H.) in an International Agency for Research on Cancer (IARC) Working Group as an author of a paper about a large cohort study on occupational formaldehyde exposure⁵ may have given that study an ‘overly important role’ when the IARC decided to upgrade formaldehyde from probably carcinogenic (group 2A) to carcinogenic (group 1)⁶. They thereby consider the IARC monograph process as highly sensitive towards unfounded opinions of individual Working Group members. However, this is not true, since IARC Working Groups consist of members selected to include all scientific views and evaluations are based on consensus or majority vote⁷. Such a statement by a group including a prominent IARC scientist leaves us puzzled.

Nevertheless, we agree with many of the arguments they laid out in their original commentary⁸, where they concluded that epidemiologists commonly overinterpret observed associations potentially leading to inappropriate public health decisions, loss of credibility, waste of resources and misguided research
objectives. Cautious interpretation of findings is important, particularly when subgroup analyses or post hoc hypotheses are involved.

However, we believe their perspective is unbalanced as it virtually ignores the deleterious effects of false-negative reporting. We all want to avoid false-positive but also false-negative findings. However, we cannot have both—it is a trade-off. Delayed recognition of true effects can have serious consequences, including some of the very same Boffetta et al.5 are concerned about. Delays of regulation of identified hazards can be harmful to the public’s health, e.g. sudden infant death syndrome (SIDS) and prone sleeping position6 or folic acid intake and neural tube defects7. Unnecessary costs can arise from endless calls for further research although the evidence appears strong enough to act8. Reducing the number of false-positive conclusions and therefore inevitably increasing the occurrence of false-negative conclusions is also not consistent with the precautionary principle, which implies that it is better to err on the side of overstating potentially false-positive results than to miss the identification of a true hazard. This principle is being applied for chemicals in the EU (Registration, Evaluation, Authorisation and Restriction of Chemicals—REACH) and is currently being considered for implementation in the USA9.

The two examples mentioned above, although involving rare outcomes, highlight the personal, financial and societal costs associated with delayed governmental action despite early signs of important epidemiologic links. With regard to folic acid intake and neural tube defects, Hibbard and Smithells10 found a strong association between a test reflecting folate metabolism in pregnant women and nervous system malformations of their babies (odds ratio >10) in a small matched case-control sample in 1965. However, routine fortification of a wide range of food products with folic acid and recommendations for folic acid supplementation were only introduced in the late 1990s after several additional studies, including randomized controlled trials, confirmed the association7. The association between sleeping position and SIDS illustrates an even more dramatic failure to recognize early signs of a true association by epidemiologists and regulators involved in translating new findings into public health action. Gilbert et al.6 estimate that over 60,000 babies have died from SIDS worldwide since 1970 due to harmful health advice because collective evidence available in 1970 was ignored, i.e. a statistically significantly 2.9-fold increased risk of SIDS for front sleeping compared with back from two pooled studies published by 1970. In fact, a false-negative finding may be worse than a false-positive because we may stop looking, while with the false-positive, the truth will eventually win out.

In contrast, the examples for false-positive findings presented in the commentary5 (acrylonitrile and lung cancer, coffee and pancreatic cancer, 1,1-dichloro-2,2-bis(p-chlorophenyl)ethylene (DDE) and breast cancer, abortion and breast cancer, genetic studies) do not show any evidence of ‘hype’, did apparently not lead to governmental or public health decisions and do not suggest the problem is common. More importantly, they highlight exactly the way science is supposed to operate. An investigator observes an interesting finding and reports it, allowing others to examine their data. Inevitably, most, if not all, false-positive results will eventually be overturned by further studies—the evolving scientific process will weed out incorrect findings. It would be odd indeed—perhaps boarding on unethical—to refrain from reporting an observed, yet controversial association, as was recently argued for cellular phone use during pregnancy and behavioral problems in offspring11.

According to Boffetta et al.5, multiple comparisons are the major cause of false-positive findings as they are often made in studies with multiple exposure metrics or multiple outcomes or when different reference categories are used. However, those comparisons, if carefully applied, can actually help to identify true hazards. Multiple exposure metrics are useful when the measure that best characterizes delivered dose is unknown2,12. Pursuing observed links of established carcinogens with cancer at sites not the primary target of their carcinogenic effect is an important strategy to discover new mechanisms. Historically, many exposure-disease associations were first reported by alert clinicians who noticed a disease excess (or deficit) in particular patients, followed by examination of possible biological mechanisms. Data on multiple diseases or causes of death can also help to address confounding. For example, in a cohort mortality study of occupational exposure to formaldehyde without individual information on smoking, in which one of us (M.H.) was involved, it could be ruled out that an observed strong association between formaldehyde exposure and myeloid leukemia was due to confounding by smoking, because no association was observed between exposure to formaldehyde and lung cancer2. In the same study, low-exposed workers were used as the reference instead of non-exposed workers because the latter may differ from exposed workers with regard to unknown confounders.

The UK Academy of Medical Sciences7 recently published a report on environmental causes of disease and how we should decide what to believe and when to take action. This report, unlike the commentary by Boffetta et al.5, appears balanced and written with epistemological modesty. It can be highly recommended.

Finally, we note that neither have we received external funding for this commentary nor have we received funding from any commercial source with interests related to the subjects discussed in this letter.
Contrary to the claim of Hauptmann and Ronckers, we did not single out any particular investigator with regard to the International Agency for Research on Cancer (IARC) working group process, either in our original Commentary or in our response to the Viewpoint of Vineis. Although criticism of the IARC Monographs was not the purpose of our Commentary or the response to Vineis, it is naive to believe that the IARC Monograph evaluation process—or similar evaluation programs, which might lead to policy, including science policy, decisions—are free of the modal conflict of interest, careerism. As noted in the Preamble to the IARC Monographs, working group members are chosen because they ‘have published significant research related to the carcinogenicity of the agents being reviewed’. Thus, although investigators participating in the decision process may come from a variety of disciplines, many have a vested interest in advancing their own research results in the deliberations, if only to increase their prestige and future funding opportunities. The routine inclusion of such self-interested researchers is a clear conflict of interest, and inhibits an open and robust evaluation of the strengths and weaknesses of the studies under examination.

A working group made up of experienced and well-trained investigators who are not professionally invested in the topic of the working group deliberations would likely provide a more scientifically open and honest evaluation of the experimental and epidemiologic evidence related to the possible carcinogenicity of the agent being examined. We doubt, for example, that many epidemiologists would agree with the characterization by Hauptmann and Ronckers of the reported association between formaldehyde exposure and myeloid leukemia in the National Cancer Institute (NCI) formaldehyde worker cohort as ‘very strong’. In view of the multiple types of