contact: Marian K. Stanley
Manager
Phthalate Esters Panel
703.741.5623

Widely Publicized Report on Study Claiming Effect of Phthalates On Male Babies Requires Rigid Scientific Scrutiny

An Interim Statement from the Phthalate Esters Panel
American Chemistry Council
June 1, 2005

A recently published study has been widely reported to claim a statistical correlation between exposure to some phthalates and structural changes in the genitalia of male infants. The report itself did not actually make that claim – its only claim is a correlation between phthalate exposure and a small anatomical change of unknown significance. The study will be thoroughly analyzed by the Phthalate Esters Panel's Toxicology Research Task Group, and a report will be posted.

Initial analyses indicate that the study has many weaknesses, suggesting that the study may not stand up under rigid scientific scrutiny. The authors themselves stated that the results need to be validated. They concede that the clinical relevance of the key index used in the study, called anogenital index (AGI), is not known. And they state clearly that none of the boys in the study had any genital defects or malformations.

A key question concerns the way the data in the study were analyzed. The statistical methods that are the basis for the paper's thesis have already been challenged. In a paper from the Statistical Assessment Service at George Mason University, mathematician Dr. Rebecca Goldin challenges whether statistically significant correlations between phthalates were indeed found by the authors, or how much "data fiddling was required to find a result." Dr. Goldin raises many questions about the study that need to be answered, and concludes that when appropriate statistical standards are applied, "not one phthalate passed the test of a statistically significant correlation." The entire article is attached below.

Questions have also been raised about the study because it disagrees so dramatically with many studies developed over decades. First, the phthalate levels found in the mothers of the measured babies were thousands of times lower than the levels that caused similar changes in rodents – and most research indicates that primates are less sensitive than rodents to phthalates, not more sensitive. For example, recent laboratory research (Kessler et al.) shows that when fed equivalent doses of one phthalate, rats and marmosets (which are primates as are humans) take up the phthalate much differently. The marmosets took up between 7.5 and 16 times less phthalate than the rats, suggesting that the chemical would be expected to have less, not more impact on humans than on rodents. In addition, the report turns some of the scientifically validated information previously gathered about phthalates on its head. One phthalate that shows few effects in rats and is generally considered very benign is said to show a correlation with AGI in this study. But the phthalate most widely known for its effects on rodents showed no correlation with AGI in this study. Finally, none of the studies purporting to show a human health effect from phthalate exposure has been validated.

"The harder the look this paper gets, the more questions it raises," said Marian Stanley, manager of the Phthalate Esters Panel. "It has also been considerably overinterpreted in the media, thanks apparently to comments made in the teleconference promoting the paper by other than the paper's authors. In all, publication of a preliminary, small study of data whose significance is not known, in a journal that is not independently peer-reviewed, needs to be treated with extreme caution until it can be properly evaluated."

Here is the paper from the Statistical Assessment Service in its entirety, reprinted with permission. Its Web address is www.stats.org.

Media Claims Phthalates (Might) Cause Genital Defects
May 27 2005
Rebecca Goldin
But the study itself did not even consider this hypothesis

The media has pulled out the scare tactics again. Armed with a recent study which will be published in the journal *Environmental Health Perspectives*, **USA Today** claimed that a "common chemical may cause defects in baby boys." This chemical substance is phthalates, found in an assortment of products, from cosmetics to fragrance to plastics. And the alleged defects are smaller genitals, improperly descended testicles, and (down the road) infertility.

In fact, the study did not examine the impact of phthalates on genitals (and none of the boys in the study had defective or malformed genitalia). The authors of the study found a correlation between what they termed the *anogenital index* of a baby boy, and the level of residual chemicals from phthalates (also called metabolites) in the mother's urine before his birth. The anogenital index is a measurement of the distance from the anus to the base of the penis, divided by the weight at the time of measurement.

Notably, the study did **not** claim there was a correlation between the level of metabolites and penile length or volume, or the size of the scrotum, despite having measured them. Perhaps they found no correlation to report.

In addition to claiming that the study was about penile size, USA Today's report is riddled with errors. The article claims there were 134 baby boys in the study, when in fact only 85 of the mothers had contributed urine. The article claims that the researchers tested four phthalates; in fact, they tested eight phthalates and found no correlation for four of them. This kind of skewed reporting only serves to undermine the honest rendering of scientific results to the public.

Correlation between AGI and metabolites: What it means and what it doesn't mean

Even if we accept that the study established a correlation between the anogenital index (AGI) and the presence of phthalate metabolites in the mother's urine, we have to ask: what does this mean?

The authors of the study draw parallels with studies on rats that are not necessarily reflective of the human response to phthalates – and this gets reflected in the media. According to Swan and coauthors, "The associations between male genital development and phthalate exposure seen here are consistent with the phthalate-related syndrome of incomplete virilization that has been reported in prenatally exposed rodents." In other words, since rats exposed to phthalates before birth have fertility problems as well as lower AGI, human boys with lower AGI will also have fertility problems.

But this reasoning is faulty. First of all, the studies on rats involve high doses of phthalates, many times higher than the human exposure rate. It is entirely plausible that a low dose of phthalates has an effect on AGI and not on fertility, while a high dose has an effect on both AGI and fertility. Without a measured effect on human male fertility, the study cannot make any conclusions at all about phthalates and fertility.

Secondly, rats and humans have different metabolic systems – an effect that can be seen in rats may not occur in humans. It is true that research on rats has been instrumental to enormous strides in biomedical research and should not be minimized; however, there are several research steps between an animal model and clinical applications. We should be cautious in assuming that what is true for rats is directly applicable to humans.

Unfortunately, the media picked up the erroneous reasoning of the authors. Most reports of the study cited the effects of large doses of phthalates on infertility in rats -- implying without a modicum of caution that our boys are going to be infertile if we don't curb phthalate exposure.

Some methodological problems

Independent of media misrepresentation of the study's results, the authors of the study have attempted to make their case stronger than it is. The most blaring problem with the study is its limited scope. While 85 boys can show a trend, these children were not nationally representative; they were drawn from three localities only.

Why does having a homogeneous sample matter? Suppose that there were a genetic disposition among a certain group of people toward a low AGI. Suppose also that this same group tends to live in areas with higher pollution and higher concentrations of phthalates in their environment. Then we might measure higher than usual phthalate concentrations, and lower than usual AGI, yet the AGI result may be entirely a result of genetics. People who live in one demographic area are more likely to have similar genetic make-ups (which could affect AGI) and also to have similar lifestyles (which could affect exposure to phthalates). The net result: with only three locations sampled, you can't tell whether the correlation found by the authors is a cause and effect, or just a coincidence.

Finally we turn to the question of statistical significance and how much data fiddling was required to find a result. Of the eight phthalates tested, four produced what the authors termed statistically significant results. However, when multiple tests of statistical correlation are used, the standard for statistical significance goes up. In other words, if the authors are attempting (and failing) to find a strong correlation and are therefore switching to a new hypothesis in order to find something, then they have to be even more convincing that the result is not occurring by chance. The gold standard of statistical significance lies in what is called the "p-value". Typically, if p is .05 or less, then one says the result is "statistically significant" and the scientific community holds it as proven. But in this case, the magic p-value should have been "p=.05/8=.006". At this level of statistical significance, not one phthalate passed the test of a statistically significant correlation between metabolites in the mother's urine and low AGI in the baby.

Another disturbing aspect of the study is the extent to which certain measurements were chosen over others. When the authors used anoscrotal distance instead of anogenital distance, they found only one of the eight phthalates had a statistically significant correlation. The babies' penile volume was measured and no mention of correlation of volume to metabolite levels is mentioned. While the authors contend that small AGI was associated to low penile volume, they did not associate it with high metabolite levels. The study gives the overall impression that the authors looked hard to find *something* in the male genital region that correlated with high metabolite levels (of at least *some* of the phthalates) in order to make a preconceived point.

Andrea Foulkes, an Assistant Professor in the Department of Biostatistics at University of Massachusetts, Amherst, contributed to this piece.