

# **1,1-Dichloro-2,2-Bis(P-Chlorophenyl) Ethylene and Polychlorinated Biphenyls and Breast Cancer: Combined Analysis of Five US Studies**

**David J. Hunter**

Harvard Medical School and School of Public Health

Thank you. I would like to thank Prof. Hirahara for the introduction and the organizers for the invitation to speak this afternoon. It has been a terrific meeting. I particularly enjoyed the session on human health.

What I will be doing is switching from the concentration on the health of the younger males and talk about one of the major causes of mortality in women in middle life, and that is breast cancer.

I will lay out very briefly the basis for the hypothesis that endocrine disrupting chemicals might have a relationship with breast cancer. I will talk about some of the general overview data, and then I will talk in a little more detail about some specific studies and the collaborative analysis that I was involved with that tried to look at the issue in some depth.

I think the basic hypothesis does not really need to be discussed in great depth here: breast cancer rates have been increasing in industrialized countries, similar to the increasing rates of testicular cancer that Prof. Lancaster mentioned. For the breast cancer rates in the parts of the U.S. that have monitored breast cancer rates very carefully, the incidence rates have been increasing about 1% a year since the middle of the 20th century.

We know that breast cancer is at least in part a hormonally determined disease and the prevailing wisdom is that a woman's lifetime exposure to estrogen is a major determinant of lifetime breast cancer risk. Since organochlorines have been implicated as endocrine disrupters in so many different systems, it makes sense to hypothesize that these chemicals could also influence the rate in women of endocrine and estrogen related diseases such as breast cancer.

The field really started when some early small epidemiological studies suggested a positive relation between some sentinel, widely used organochlorine chemicals: DDE is the major metabolite of DDT, and PCBs.

These studies, if we took a snapshot of the data as of about 1993, the studies were fairly small. At that point there were six studies, which had been case-control studies. They all had less than 45 cases, so 45 cases and 45 controls. One of them was as small as eight cases. So they were very small by epidemiologic standards.

As you look at these studies you get the impression that overall the levels of DDE were higher in the adipose tissue from cases compared with the controls, and the same was true for levels of total PCBs.

In addition to their small size, most of these studies did not meet the standard criteria for best epidemiological study design, so they really could be only hypothesis-generating at this point. What really I think got the field going, and to some extent contributed to putting the endocrine disrupter hypothesis on the map at least in North America was the publication of the first prospective study, published by Mary Wolf and colleagues from the New York City women's health study.

They looked at 58 cases of breast cancer occurring after a blood sample was taken and 171 controls. They compared the top versus bottom decile of exposure for these compounds, i.e., looking at the top 10% of the distribution compared with the bottom 10%. They saw a significant elevation in risk associated with serum DDE (the relative risk was about four-fold) and an elevation with total PCBs, the elevation was about two-fold. This was not actually formally statistically significant.

This really got a lot of interest in the field. Breast cancer is such an enigmatic disease in many ways, and this really did suggest that one of the major potential causes could be related to environmental estrogens.

A lot of studies have been done since, and I am going to briefly summarize them.

In many ways, the preferred tissue for analyzing organochlorines would be adipose tissue. The problem is it is very difficult to get that in anything other than a case-control context: you have to stick a needle under the skin and get some adipose tissue.

It actually is not nearly as invasive as it might sound. It is really no worse than having a blood sample taken in most hands, but it means that there is no prospective data because nobody has been able to do this to several 10,000s of women and store the samples and be able to go back. So all of these are case-control studies, but they are large and in many cases well done case-control studies.

They have all been essentially null with respect to DDE in adipose tissue. A large European study, in fact, observed a statistically significant inverse trend: the higher the level of DDE, the lower the risk of breast cancer. Some large, more recent U.S. studies have essentially observed no association. A large study in Canada similarly saw no association.

So these studies do DDE in adipose tissue certainly are not indicating any positive association between DDE, the major metabolite of DDT, and risk of breast cancer.

For PCBs it is a little bit more mixed. The big European study has not looked at PCBs or reported on PCBs yet. Some of the North American studies, a large study in Connecticut really saw absolutely no association with total PCBs and adipose, or looking at groups of congeners broken out separately from total PCB score.

A study in New Jersey and New York State did not really see much overall but did see a positive association with congener-183 but not with the others. The Canadian study saw a positive association with two of the congeners for pre-menopausal women, but two different congeners for postmenopausal women.

I think the overall summary is that there is really no good evidence that the total PCB burden is associated with risk of breast cancer in these studies using adipose tissue. There is some question about congener specific analyses, but no uniformity across the small number of studies.

Turning to blood levels, the advantage with blood levels is that they are a little easier to obtain, by and large, than getting an adipose tissue sample, and also we can tap into the prospective serum and plasma banks that have been built up and so we can do larger studies more quickly and we can use a prospective design if the blood banks have been set up.

This, essentially, are all the data I could find, arrayed. I will show a number of these slides, so I will take you through this. This is arrayed left to right in order of the strength of the association. The squares indicate the relative risk between the highest level of exposure as defined in study and the lowest level of exposure. The error bars represent the 95% confidence intervals, and the line at 1 is the point at which the relative risk is 1; in other words, there is absolutely no association comparing high with low categories.

The relative risks above 1 suggest a positive association; relative risks below 1 suggest an inverse, potentially protective association. The study that got everybody excited was Mary Wolf's study, and it is actually the study with the second highest relative risk after all of these studies have come in.

Shortly after that, Nancy Krieger, analyzing data from the Kaiser Permanente blood cohort from the San Francisco Bay Area, looked at breast cancer and broke it up by ethnicity. She looked at 50 incident cases among black women, 50 incident cases among white women, and 50 incident cases among Asian women, and saw non-significant increases for the white and black women but no association, even a little inverse, for the Asian women. Overall, combining those three groups, the result was statistically not significant.

So, I think it is fair to say that the bulk of that data hovers around 1; there are a few positive studies, there are a few inverse studies, but no real consistency related to a positive association.

Now, those were studies in North America, and it is obviously interesting to look elsewhere in the world, partly because, maybe not in Denmark, but partly in some other countries, such as Mexico, for instance, DDT is still being used in certain areas, and so body burdens might be higher. By looking in countries where the body burden is higher we might see an association that was not apparent in North America, where the body burdens tend to be lower: DDT was banned in the early 1970s. I think it is fair to say there is really not enough data to reach any conclusions.

Here are two studies in Mexico, very similar study designs and it is not obvious why one of them saw a positive association and the other was really very null. A study from Columbia saw a positive association; a study from Brazil was essentially null.

I think this is one gap in the science here: studies from developing countries where DDT is still being used, where body burdens may be higher, where breast cancer rates that tended to be very low historically are starting to climb. There is probably some more data that we need from those studies before we can really conclude what is happening around the world.

Here are studies for the total PCB measures as a sum. Mary Wolf's study from New York City is now the highest of all these studies that were available at the time. But in most cases, subsequent studies have come in, some of them weakly positive, some of them weakly inverse, nothing with any sort of strong overall relation. Arraying the data this way, you can get a visual impression that it is not a strict quantitative estimate of the best summary estimate.

What I would like to do is to go into a couple of studies that give more quantitative estimates and allow us to look within interesting subgroups. I will talk, first of all, about the Nurses' Health Study, which is a cohort study that I work on. In this study, the parent study has about 121,000 women under follow-up since 1976.

In 1989 to 1990 almost 33,000 of those women gave a blood sample, which was archived. In these analyses, we looked at 381 cases of breast cancer which occurred before June of 1994, in the four or five years after the blood was donated, and matched 381 controls on age, by single year, and by a number of other risk factors: menopausal status, post-menopausal hormone use, day of blood draw, timing of blood draw, etc.

We were concerned to try and make the data as comparable as possible with Mary Wolf's study, so her lab actually analyzed the data by gas chromatography. We buried in the samples blinded split replicate samples, so she did not know that these quality control samples were in the batches. The median correlation of variation was 5% for DDE and about 12% for the total PCB measures, which I think is very respectable for these difficult to measure low-concentration compounds.

We lipid-adjusted — it did not really matter whether we lipid-adjusted or not. Looking at the data, let us just look at this column: the median values for the lipid-adjusted data for cases were actually a little lower than for controls, so in the opposite direction of there being a positive association, and were essentially equivalent for the PCBs.

When we looked a little further, breaking the data up into five categories of increasing exposure, quintiles, 20% of the data, this is the low, this is the highest, and this is the \*role reverse\* in confidence interval associated with each category relative to the lowest category, there is no evidence of any positive association; if anything it is a little bit inverse or perhaps even u-shaped. There is not a significant inverse linear trend, and that is compatible with there really being no association of plasma DDE with the risk of breast cancer.

Looking at PCBs: a similar picture. The relative risks were all below 1, not above 1. Everything is non-significant, no evidence of a positive or inverse association. All these data are adjusted for body mass index and other breast cancer risk factors.

Now obviously, lumping all the PCBs together into one score, even though that was what was done in most of the original studies, is not the most biologically motivated way of doing it. We know that the congeners can have very different specific biologic actions. We can have a little bit of look at this by looking at the four most common congeners one by one, or adjusting for the other; whatever we did, it did not really make much difference: no positive association, no significant inverse association. Looking at the congener specific analyses was pretty compatible with the overall lack of any association of the total score.

Epidemiologists always like to break their data into subgroups. Sometimes this is motivated by a prior hypothesis for a particular subgroup, and sometimes it is just good analytic practice to see if you can thoroughly examine the data to see if there are subgroups in which there may be associations. Any good epidemiologist will always come up with a hypothesis for any subgroup finding that one sees.

It is not too hard to justify, I think, why you would be interested in looking at these associations within body mass index or obesity. These are lipophilic, and that is going to have some association with body burden and blood concentration.

Here, the data are broken into three categories. These are relatively lean women, these are women who meet the WHO definition of obesity, and these are women who meet the WHO definition of overweight. This is a North American study, so more than half the women are in these two categories.

This is now looking at tertiles: we have got three categories instead of the previous five categories because we have broken the data down and we are trying to get stable estimates within these strata. Everything is compared with the lowest category, the lowest tertile of exposure.

There is actually a statistically significant interaction. In other words, statistically the slope here is different from the slope here and the slope here, with perhaps a hint of a positive association among lean women and actually a statistically significant inverse association among obese women for high versus low tertiles of PCBs. We could get going and make a hypothesis about why PCBs might be protective among obese women, might be a weak risk factor among lean women.

We also looked at parity and lactation. For the persistent organochlorines pollutants, one of the best ways of a woman reducing her body burden is to lactate and pass a lot of that out in the breast milk. Again, you could make a hypothesis that this association might be different for nulliparous women, women who have never had a pregnancy. It might be different for women who had had a pregnancy between those who did not breast-feed and those who did.

Here in the very small group — about 11% of the women who have not had a full-term pregnancy — there appeared to be a positive association, marginally significant, it actually was statistically significant. Comparing high versus low PCBs in this small group of nulliparous women, and really nothing is going on in the other stratum. Again, this is actually a technically significant interaction.

One of the problems with interaction analysis in the literature is that it is a rare interaction that actually makes it in this form in a table or a graph into the ultimate study publication. Editors do not like you to fill up their pages with page after page of null tables.

So the literature tends to be very biased toward positive interactions or significant interactions or interesting interactions, and if you are lucky you will find somewhere in the results or discussion a brief mention of all the interactions that were looked at that might not have been null. But you do not have any data to work with quantitatively out of the literature if you have just got that single statement.

So we were concerned to put the best estimates overall on the large recent studies and also see if these subgroups behaved consistently across studies by doing a parallel analysis of five large U.S. studies: a case control study conducted in western New York State; the Nurses' Health Study I talked about; a

long-running prospective study from Johns Hopkins, the Washington County study; and two more recent case-control studies, one from New York City and one from Connecticut.

These are the major investigators involved, but there were a much larger number of collaborators in each of these studies I would like to acknowledge who contributed to the data.

The reason that these studies came together was somewhat arbitrary, but stimulated by the real interest in this field, particularly after the initial positive prospective study. We were funded, as a group of studies in the northeast U.S. to examine this hypothesis. One of the criteria for funding was that all the studies had to have good data to control for potentially confounding factors to look at potential interacting factors. All of these studies together give a database of 1,400 cases and 1,642 controls.

Briefly, the statistical methods: we analyzed each study using study-specific categories. We controlled for confounding in a standardized manner, controlling for the standard classical breast cancer risk factors, and other factors that might influence body burden of DDE and PCBs. We planned analyses of effect modification or interaction by these covariants, and we developed the pooled estimates using a random effects model.

The individual studies: (r stands for retrospective) these are case control studies. I have talked about the Harvard Nurses' Health Study. The Johns Hopkins prospective study was particularly important because they drew their blood early in 1974, and this was just at the time that DDT was being banned in the U.S. So this was really only the study that had a measure of DDE that applied to the time when DDT was still in reasonably widespread use, or at least commercially available in the U.S.

Each of the labs used different methods except the Harvard study and the Sinai study as I mentioned before. This is a problem, and all of the labs used different methods, so they measured different numbers of congeners. They all measured the major congeners that contribute 80% or more to the total, and they did participate in a laboratory round-robin and got quite respectable results looking at blended specimens that were sent around to each of the labs, suggesting that at least within the labs the results are quite accurate.

Very briefly, this is the distribution of lipid-adjusted DDE in the studies. These are the percentiles: 20th up to 80th percentile. I just want to point out that the Hopkins' study had the highest values, consistent with it having the oldest blood samples, and the Yale study and Mt. Sinai studies have the lowest values, consistent with those being the recent studies.

So, it is very compatible with the secular decline in DDE concentrations in adipose and plasma in adults in North America. There is less variability in the PCB measures.

Here are the data. These are the estimates for the top versus bottom category for DDE in breast cancer. None of the studies saw any significant positive association, maybe a hint of something here, otherwise the overall estimate is essentially right on 1. This is the pooled estimate with its \*confidence interval\*; relative risk is about 1, with a very tight \*confidence level\*. For PCBs, it is pretty much the same picture — none of the studies were statistically significant and the overall estimate was also almost exactly 1 — completely null.

Looking at the subgroups: this is the subgroup for nulliparous women, where we saw the positive association. That disappears off the graph here, but none of the other studies did, and the overall estimate is null.

The Buffalo group had been quite excited about this: they had a hypothesis why PCBs would be more risky in women who did not breast feed, but none of the other studies were able to replicate that and the stratum specific estimates are almost exactly 1.

Looking at body mass index, in our study here we have a significant inverse association for the overweight or obese women, and in fact actually so did a number of the other studies, but not all. The

overall estimate is not statistically significant. We saw the hint of a positive association in lean women, but the other studies did not and that is not statistically significant.

I think this is a nice example of the way with a collaborative analysis and planning analyses that it is possible to see if subgroup findings replicate and get that into the literature in a fairly standardized way and as with so many subgroup analyses, they did not replicate once we were able to look across all studies.

So in summary, I think there is very little support for the idea that there is a positive association between plasma or serum concentrations of DDE or PCBs with breast cancer risk, and the combined analysis was also unable to support findings that had been reported from the individual studies for some of these covariants limited to subgroups of the populations. I think that as a general point that this sort of multi-center study, preplanned analysis — even if each of the studies is essentially operating independently — can be very useful.

But I think it is very important to note what we cannot conclude here, limitations of the data: we only address DDE and PCBs and by no means did we address the whole length and breadth of the endocrine disrupting chemical hypothesis. We do a reasonable test of the most common and widely spread classes of compounds, but by no means all of the myriad compounds, some of which have been talked about at this meeting.

Perhaps the most serious problem is that studies in women in middle life do not address a lot of what we have been talking about with respect to the subsequent risk of chronic disease, particularly cancers. If endocrine disrupting chemicals *in utero*, in childhood, during puberty, or in early adult life influence a woman's risk of breast cancer decades later, these data do not necessarily address that because even though the half-life of these compounds is long, clearly those data do not tell us anything about a woman's *in utero* or childhood exposure.

I think this is going to be a very, very difficult issue because probably the only study designs that will actually give us an interpretable answer are if we follow Frank's 8,000 women for 50 years and some of the other birth-cohorts that are already ongoing, it will literally take 50 years or 60 years before women get into the breast cancer age group, and this specific hypothesis could be addressed. Thank you.

## Q&A

Hirahara: Thank you very much, Prof. Hunter. We would like to open the floor for discussion. Are there any questions?

Q: Thank you for that study. I realize it started right after the hypothesis came out. What you did was actually prove that DDE and most PCBs are probably not estrogens, which is really good. If you had actually found a relationship, I think a lot of us would be surprised, given the fact that breast cancer seems to be tied very tightly with exposure to estrogens.

I wonder if, given all of the energy that you put into this, if you are going to follow that up by actually looking at OCs and other compounds which we know are estrogens, at least are able to interact with the human estrogen receptor and actually turn on genes.

Hunter: So the point you are making basically is, we look at one class of chemicals; what about all the other classes of chemicals, and as the science moves on, we know more about the specific environmental chemicals that have higher estrogen potency.

The answer is that I think a lot of people would like to follow up. The major limitations are how you measure those other compounds. Many of the measures currently require a very large amount of blood or an adipose tissue sample that is larger than any of the current studies have, although some of the studies are setting up to look at dioxin, for instance, taking 100mL of blood to try and look at that specifically.

They will be case-control study designed, and so they will have to be designed as well as possible. It will be a long time before we have got any prospective data.

Somebody called epidemiology “the art of the possible.” I think we have really gone about as far as it is possible with the prevalent human studies and databases we have.

It is going to be a real challenge to explore it into different areas unless the technology allows

us to measure some of those compounds at the  $\mu\text{L}$  level, which is basically the sort of blood sample we have, plus we have got all the problems of half-life, etc., that many of those compounds have much shorter half-lives and you really need an integrated exposure over more time than a single blood measure will give you.

Q: One study already has demonstrated that dieldrin seems to have a fairly strong relationship with breast cancer in Europe. I think that there are a number of organochlorines that have clear interactions with the human estrogen receptor.

Certainly the technology, I know at the CDC and several other institutes around the world, now have the ability to measure those OCs literally in drops of blood.

So I think that is someplace where, if we are thinking about going forward, I do not know if we really need to keep hammering DDE, which we know is an anti-androgen, not an estrogen in humans, but PCBs for the most part do not interact with the estrogen receptor, unless we start doing more sophisticated analysis and look at the hydroxylated and sulfonyl forms, other forms of PCB that we know actually do interact.

But you are correct. They are short half-life, there are very rapid, and very difficult to measure. I think it is an important point that we have to address.

Hunter: It is, and so the history here has been a positive study in null studies. The western New York Study, for instance, looked at dieldrin and saw absolutely no relation, so they were not able to confirm the Danish study result.

Hirahara: Any other questions? OK, thank you very much, Prof. Hunter.