Endocrine Disruption - The Trouble with Mixtures

Andreas Kortenkamp
University of London, UK

The levels of some individual endocrine disrupters found in the environment and the tissues of organisms are usually very low. In order to observe effects in laboratory assays, far higher concentrations need to be administered. This one factor complicates enormously the discussion about endocrine disruption and possible health risks to humans and wildlife. An example with four common pesticides should serve to highlight the problem. The range of concentrations of p,p'-DDE, β-HCH, p,p'-DDT and o,p'-DDT found in human serum in a variety of countries is several orders of magnitude lower than the no-observed effect levels of these chemicals in, for example, the E-screen assay. This enormous discrepancy between environmental concentrations and effective levels in laboratory assays has led many researchers to believe that synergisms between these chemicals need to be invoked to explain possible effects in humans and wildlife.

The alternative position is that adverse effects from estrogenic endocrine disrupting chemicals are very unlikely to arise because their potency is low when compared to the natural hormone estradiol. Thus, the endogenous hormone is assumed to be so potent that enormous doses of endocrine disrupting chemicals are required to make a difference. This idea is often used to dismiss claims of possible adverse health effects of EDCs.

Here, I will present data and details of experiments that our group in London have designed to respond to these two issues. But before we can proceed with this, I will need to briefly clarify a number of issues relating to mixture effects and how they can be assessed.

Mixture effects are usually classified in relation to the effects that are expected to occur on the basis of knowledge about the potency of individual mixture components. If the observed effects of a mixture are larger than the expected ones, we can call the overall mixture effect synergistic; if they are smaller than expected, there is thought to be antagonism; and if expectations are met, the agents are thought to act in an additive way. One issue that has complicated discussions in the mixture field for a long time is to decide what a reasonable and well-founded additivity expectation should be, and more importantly, how it can be calculated quantitatively.

There are three ways of calculating additivity: effect summation, independent action, and concentration addition. Each of these concepts was developed independently to suit special purposes and experimental conditions. For example, the concept of independent action is frequently used for the assessment of combinations of anti-cancer drugs, whereas concentration addition has been employed for mixtures of environmental pollutants. I will briefly take each concept in turn and list its underlying pharmacological assumptions.

In effect summation, the effects of individual mixture components are simply added up. It is an intuitively appealing and easy to use method, yet it is frequently overlooked that this concept can only be applied to linear dose response relationships. In all other cases, that is, when curves are of the familiar sigmoidal shape, effect summation is unreliable. We will illustrate this point later in the talk.

Independent action, as the name suggests, assumes that all mixture components act independently of one another. The concept is usually applied to mixtures that consist of chemicals with diverse modes of action or different sites of action.

Concentration addition rationalizes the alternative scenario. It assumes that all mixture components interact with the same biological target, essentially in a similar fashion. To calculate the additivity, or perhaps better, zero-interaction expectation in this concept, doses that produce the same predetermined

effect are added up. Thus, the name concentration- or dose-additivity.

The starting point of our work with mixtures of estrogenic endocrine disrupting chemicals has been to ask whether the effects of multi-component mixtures can be predicted accurately on the basis of information about their individual potency. Because there have been acrimonious disputes about the general applicability of independent action and concentration addition, we have compared the performance of both concepts. In view of its popularity and widespread use we have also included effect summation.

We have chosen the yeast estrogen screen as our model to assess the predictability of combined effects of estrogenic chemicals. As can be imagined, work with multi-component mixtures places very high demands on the chosen assay in terms of reproducibility and low variability. The yeast estrogen screen fits these demands; it is reliable and very reproducible.

Figure 1 below shows the regression models for our chosen mixture components. They include a wide spectrum of estrogenic chemicals such as hydroxylated polychlorinated biphenyls, bisphenol A, genistein, benzophenone, resorcinol and others. On the very left, the curve for estradiol is shown, for purposes of reference.

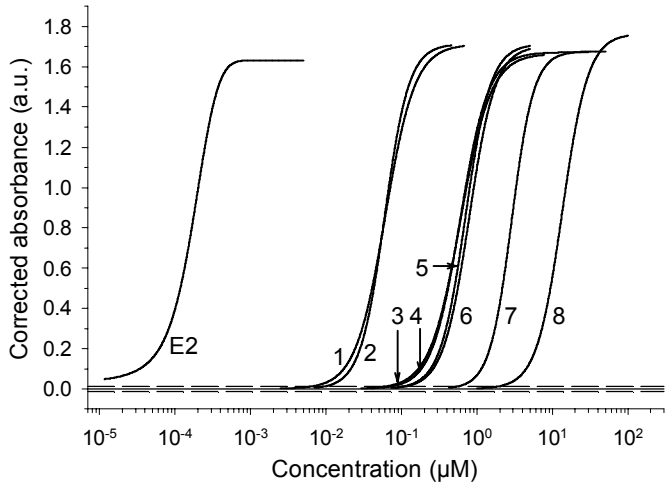

Figure 1: Regression models for estrogenic chemicals 1 - 8 in the yeast estrogen screen. E2: 17beta estradiol. Data from Silva et al. 2002

The concentration response data shown in Figure 1 were used to calculate the expected effects of a mixture of 8 chemicals with a mixture ratio in proportion to their potency. We have calculated entire dose response curves for this mixture by using effect summation, independent action, and concentration addition (Figure 2). The curves for independent action and effect summation were quite similar in the low-dose range. In contrast, concentration addition (red curve), predicted effects at much lower concentrations. We, then, tested these predictions experimentally. The observed mixture effects (circles) agreed excellently with concentration addition and this was true over the entire range of effects.

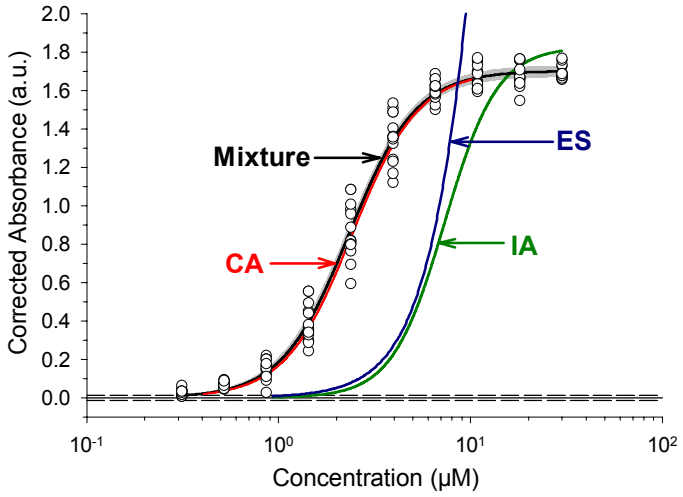

Figure 2: Predicted and observed mixture effects for a mixture of 8 estrogenic chemicals. The predictions were made for a mixture ratio in proportion to the potency of individual mixture components, based on the single-agent data shown in Figure 1. CA: concentration addition prediction (red), IA: independent action prediction (green), ES: effect summation prediction (blue). Data from Silva et al. 2002.

The message from these experiments is that the effects of quite complicated mixtures can be predicted accurately on the basis of data about their individual effects, provided the correct prediction or assessment model is used. Had we used effect summation, or independent action as our prediction model, we would have concluded erroneously that the combined effect of these 8 chemicals would have been synergistic. This is because the observed mixture responses were higher than those effects predicted by these curves. Instead, we concluded that the mixture effect in this case is clearly additive and that concentration addition is the appropriate model because all chemicals act in a similar fashion by activating the estrogen receptor.

Quite understandably, there is much excitement with a valid demonstration of synergistic mixture effects, and this has triggered much controversy in the endocrine disrupter field in the past. Certainly, synergisms will always heighten concerns. However, we believe that the question of whether there are synergisms or not, is somewhat misleading, if considered in isolation. Equally relevant, we think, is the issue of whether there are combination effects even when each individual mixture component is present at levels below their individual effect threshold.

Here, the two concepts of independent action and concentration addition predict quite different outcomes. According to independent action, there should be no overall combined effect if all individual components produce zero-effects on their own.

In contrast, concentration addition predicts that there will be a joint affect even when all components are present at sub-threshold concentrations, provided there is a sufficiently large number of chemicals. The implications of these two scenarios for real existing mixtures in environmental media are perhaps only too obvious.

A key issue in addressing the problem of mixture effects at very low doses is: what is zero-effect? If we consider the mathematical formula for independent action, it becomes obvious that the combined effect of

100 chemicals is only really 0 when the individual effects of all components are 0, too.

If, in fact, they are only slightly higher, the joint effects can be dramatic. Thus, 100 agents each producing only 1% of a theoretically maximal effect will yield a combination effect of 63%. Even if they only induce 0.1% of an effect individually, the joint effect will still be 9.5%: very different from 0.

A fundamental difficulty arises: Because it is impossible to distinguish very small statistically insignificant, albeit real existing, effects from background noise, zero-effect-levels cannot be measured reliably. For these reasons, toxicology has introduced the idea of no-observed effect levels (NOEL), or no-observed effect concentrations (NOEC). NOELs and NOECs are intended to be approximations of zero-effect levels. They are defined as the highest tested dose or concentration that does not induce effects statistically significantly different from those seen in controls. How reliable are NOELs and NOECs in the context of mixture testing?

We have compared the NOEC estimated for o,p-DDT in the E-Screen assay with the effects predicted for the NOEC by using the regression model fitted to our data. The regression model estimated an effect of 0.3 for the no-observed effect concentration, which, given that o,p'-DDT induced the maximal proliferative effect of 3.2, is 9% of the maximal effect; far greater than zero!

The problem we see here is appreciated in the ecotoxicological literature. It is now well established that NOECs tend to be higher; the fewer the number of tested concentrations was, the higher the biological variability of the assay system, and generally the poorer the overall data quality.

Far from being a valid approximation of zero-effect-levels, it has to be concluded that NOECs define boundaries within which the occurrence of effects can neither be confirmed nor ruled out with confidence, and that is something very different from zero-effect levels. We had to conclude that NOECs are of limited use when designing mixture experiments that address the question as to whether there are combination effects at very low doses of the individual mixture components.

Effect data estimated on the basis of regression models, so-called benchmark concentrations, are increasingly seen as alternatives to NOECs. EC1 levels, i.e. concentrations that produce only 1% of a maximally possible effect, are regarded as a good alternative standard. Let me emphasize that EC1s cannot be measured directly. In many assays, it is not even possible to measure reliably effects that represent 10% of a maximal response.

Further, we had to avoid designing trivial experiments. In our case, this would have happened had we combined too few chemicals at very low doses, such that the combined effect would have been too low to be measurable. Thus, how many chemicals have to be combined so that combined effects at levels around or below their respective EC1 can be observed?

The answer depends not least on the steepness of the dose response curve of individual chemicals in the mixture. The German mathematicians Boedeker and Drescher have proposed a solution to this problem. They found that the shallower the slopes, the more chemicals have to be combined to observe mixture effects. Given that the slope of many estrogenic chemicals in the yeast estrogen screen is between 1 and 2, we chose a mixture of 8 chemicals.

We proceeded in the following way: using the concentration response relationships of the individual 8 chemicals (see Figure 1), we estimated EC1. We divided this by 2, and mixed all chemicals at these concentrations, that is, 50% of their EC1. Would there be an observable mixture effect?

Our experimental results (Figure 3) demonstrate that the answer to this question is a decisive yes. Depicted on the left of this bar chart are the responses expected to occur from each individual of the 8 components in our mixture. They are too low to be measured.

The cross-hatched bar labelled ES shows what would be expected had we simply added up these individual responses. The larger hatched bar labelled CA is the concentration addition expectation. The black bar (MIX) shows the experimentally observed effects, in excellent agreement with the concentration

addition prediction.

It is important to note that effect summation dramatically underestimated the observed effect. Again, had we used this concept, in order to assess the results we would have concluded erroneously that the mixture effect is synergistic.

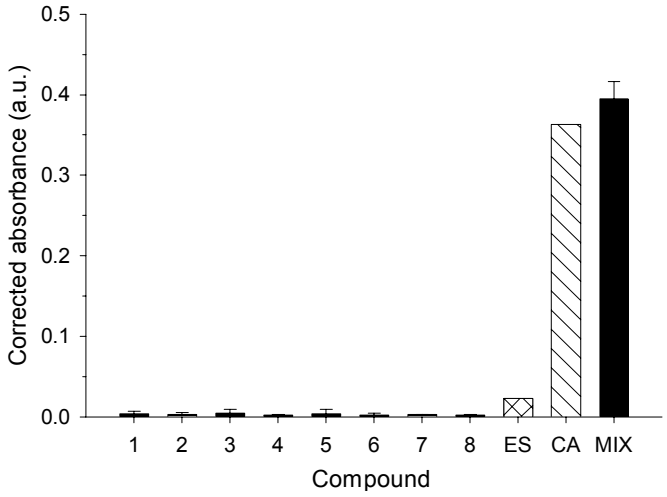

Figure 3: Mixture effects at low concentrations of the individual mixture components. Data from Silva et al. 2002.

This experiment perhaps represents a regulator's nightmare. It shows the pitfalls of the current risk assessment paradigm with its exclusive focus on single chemicals, that largely ignores mixture effects. We demonstrate that it is quite possible that an individually acceptable dose of one chemical combined with an acceptable dose of a second and so on can lead to significant mixture effects.

We became interested in addressing the second point; would weakly estrogenic chemicals be able to modulate the effect of potent steroid hormones? As mentioned earlier, this idea is often used to dismiss possible concerns resulting from exposure to xenoestrogens.

Again, we chose the yeast estrogen screen to approach this problem. We extended the number of chemicals in our mixture from 8-12 and included estradiol.

Next, we mixed all 11 weakly estrogenic chemicals in proportion to their potency, again to ensure that not one single chemical dominated the joint effect. This xenoestrogen pool in turn was then combined with estradiol at a ratio of one estradiol to 50,000 of the pool of xenoestrogens. We calculated the predicted concentration response curves assuming additive effect by using concentration addition. We also included effect summation. The results showed that the combined effect of this mixture was additive, with excellent agreement with the prediction made by using concentration addition.

To highlight the impact which the weak xenoestrogens produced on the effect of the hormone, we chose to depict individual and joint effects of a mixture at a concentration of 5 micromolar of the total mixture. This concentration produced a half-maximal effect in the yeast estrogen screen. The bar chart in Figure 4 shows that the concentration of estradiol alone, shown on the left with the black bar, gave an effect of approximately 0.4.

The levels of all other chemicals in our mixture were too low to be measured directly, yet when combined with estradiol, they led to almost a doubling of the individual effect of the hormone (grey bar

labelled MIX), well in agreement with the concentration addition prediction which is shown in the hatched bar labelled CA. The white bar shows effect summation.

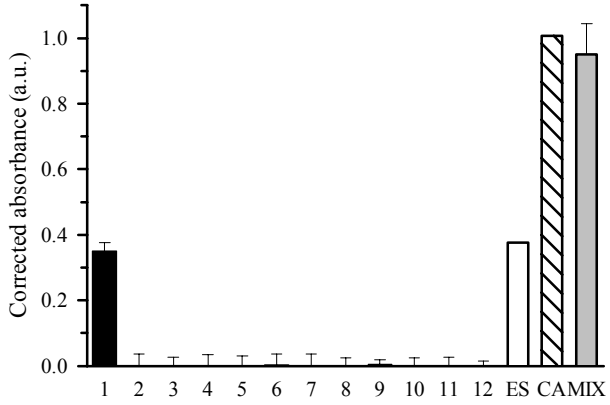

Figure 4: The impact of weak xenoestrogens on the actions of estradiol. Data are from Rajapakse et al. 2002.

Our first conclusion is a reassuring one. Given the tendency to ignore the assessment of mixture effects because of its perceived difficulty, we can say: it can be done. It is possible to predict the effects of quite large multicomponent mixtures with surprising accuracy.

The second conclusion is an unsettling one; the fact that single chemicals are present at low, ineffective concentrations cannot be taken to signal absence of risk. Depending on how many similarly acting chemicals are also present, the assumption of no hazard can be plain wrong. This is especially virulent in the context of endocrine disrupting chemicals, where we do not know with certainty how many chemicals are relevant in the environment and in humans.

Thirdly, the perceived weakness of man-made estrogenic chemicals cannot be used to rule out possible impacts on the effects of endogenous steroid hormones. If sufficiently large numbers of chemicals are combined, substantial modulations of hormone effect can occur, perhaps perturbing physiological equilibria.

To achieve the goals of our study, we had to rely on high-throughput *in vitro* assays. In this way, we were able to handle reproducibility and cost. It remains to be seen whether our methodology can be applied productively to *in vivo* assay systems. The variability of *in vivo* assays is considerably higher than that of *in vitro* assays. Another major challenge is to explore the utilization of analytical mixture approaches in epidemiological studies.

Acknowledgements

Jo Payne who was supported by the Biochemical and Biological Research Council in the United Kingdom carried out the E-screen assays. Nissanka Rajapakse did the large 12-component mixture experiment, he is grateful for support from the School of Pharmacy London for a PhD studentship. Elisabete Silva carried out the 8-component mixture. She was supported by a grant from FCT, Portuguese Government. We gratefully received the yeast cells from Prof John Sumpter, Univeristy of Brunel, United Kingdom. I thank Rolf Altenburger, Martin Scholze and Thomas Backhaus from the University of Bremen

in Germany for many stimulating discussions. Finally, I would like to acknowledge financial support from the European Union (EVK1-CT-2001-00100, QLK4-CT-2002-00603).

References

Rajapakse N, Silva E, Kortenkamp A (2002)
Combining xenoestrogens at levels below individual No-observed-effect concentrations dramatically enhances steroid hormone action
ENVIRON HEALTH PERSP 110 (9): 917-921 SEP 2002

Silva E, Rajapakse N, Kortenkamp A (2002)
Something from "nothing" - Eight weak estrogenic chemicals combined at concentrations below NOECs produce significant mixture effects
ENVIRON SCI TECHNOL 36 (8): 1751-1756 APR 15 2002

Q&A

Morita: Thank you very much, Dr. Kortenkamp. Very interesting and important work on the combined effects of multicomponent environmental hormones, especially the very beautiful prediction of multicomponent mixtures is very impressive. I would like to invite questions or answers or comments. Yes, please.

Q: As the chairman said, I also appreciate your beautiful presentation because I have been speculating even if synergistic pattern of prediction is additive. One thing I was puzzled in your presentation was that you use the term synergistic. To my knowledge and understanding, we use the word synergistic to mean over the estimation of additive. What is your opinion on this?

Kortenkamp: The definition of synergism we have used is in line with what is discussed in the specialist mixture literature. To simplify matters, synergies are always defined as effects that exceed those you expected. There has been a lot of confusion in the specialist literature relating to terminology. Some people call synergistic effects potentiation or supra-additivity and so on and so forth, but we have stuck to this fairly straightforward definition.

Q: So there is some confusion there.

Kortenkamp: Yes.

Lamb: Jim Lamb, BBL Sciences. That was a very nice talk and nicely presented. I do not know if you are familiar with the work being done in the United States on benchmark dose and cumulative risk assessment in pesticides. In that case they are using an ED10 for cholinesterase inhibiting pesticides.

There is a discussion in the case of individual pesticides they are using a lower confidence level of the ED10. In the case of cumulative risk assessment they are using the central tendency and then adding them together much like you presented today to come up with a risk assessment.

One, I was wondering on your estimate of the EC1, is it a central tendency or is it a lower confidence limit, and which do you think more appropriate in this adding of different substances which follow the same mechanism of action?

Kortenkamp: We have used a slightly different approach. We have based our EC1 estimation on the best fit regression models we have seen and then just read it off and projected it onto the concentration axis. This allows you also, of course, to estimate the confidence intervals, if you take into consideration what the confidence limits of your best-fit regression model is. But this is what we have used.

This is a quite exacting standard, and we did this deliberately because in most cases, as you know, you cannot measure EC1s directly, not even EC10s in some assays. I believe in the context of mixture testing this complication will come to the fore inevitably, because when you sometimes do not see anything this does not necessarily mean it is not there.

Lamb: Of course. Thank you.

Morita: Last question, yes.

Sekizawa: Sekizawa, from the National Institute, Health Sciences. I think you know in country already international body we use the group ADI approach to estimate risk from the chemicals which have common metabolites. But in this case, especially on the endocrine disrupters, I can agree that you may obtain additivity data from a simple *in vitro* assay, like a yeast screen.

But if you considered a whole body when you have feedback control or some other effect such as shown by Dr. vom Saal that if you have phytoestrogen then you may have different body weights, then you must be more cautious

about additivity or multiexposure. You may not predict from *in vitro* assay what will happen in the whole body. That is my comment.

Kortenkamp: Absolutely, I could not agree more with you. Absolutely. This is in fact not what we are trying to do. You can never, ever change from one assay to the other. You just have to do the experiments again.

We are currently involved with a large project where we are going to test whether the approach we have used with yeast can be productively used in animal experiments. We want to see whether the predictability of combined effects is equally good.

Secondly, I am very grateful you mentioned *in vivo* assays. It may well be that *in vivo* we will observe some synergism. I am not saying we will, but it cannot be ruled out, because the basis for many known synergisms are interactions at the toxicokinetic level and that is something we cannot measure and model with simple *in vitro* assays.

So, *in vivo* is very important from that point of view, but I would not hazard any guesses as to outcome. The bottom line is always additivity. What I would like to emphasise strongly, however, is that everybody focuses so much on synergisms in the endocrine disrupter area; what we are saying is not to forget additivity. Additivity, practically speaking, may be of great relevance.

Morita: Thank you very much.