



Review article

Ten years of research on synergisms and antagonisms in chemical mixtures: A systematic review and quantitative reappraisal of mixture studies

Olwenn Martin^{a,1}, Martin Scholze^{a,1}, Sibylle Ermler^a, Joanne McPhie^a, Stephanie K. Bopp^b, Aude Kienzler^b, Nikolaos Parissis^b, Andreas Kortenkamp^{a,*}

^a Brunel University London, Institute of Environment, Health and Societies, Kingston Lane, Uxbridge, Middlesex UB8 3PH, United Kingdom

^b European Commission, Joint Research Centre (JRC), Ispra, Italy

ARTICLE INFO

Handling Editor: Paul Whaley

Keywords:

Chemical mixtures
Synergism
Antagonism
Mixture
Interaction
Dose addition
Independent action
Chemical risk assessment

ABSTRACT

Background: Several reviews of synergisms and antagonisms in chemical mixtures have concluded that synergisms are relatively rare. However, these reviews focused on mixtures composed of specific groups of chemicals, such as pesticides or metals and on toxicity endpoints mostly relevant to ecotoxicology. Doubts remain whether these findings can be generalised. A systematic review not restricted to specific chemical mixtures and including mammalian and human toxicity endpoints is missing.

Objectives: We conducted a systematic review and quantitative reappraisal of 10 years' of experimental mixture studies to investigate the frequency and reliability of evaluations of mixture effects as synergistic or antagonistic. Unlike previous reviews, we did not limit our efforts to certain groups of chemicals or specific toxicity outcomes and covered mixture studies relevant to ecotoxicology and human/mammalian toxicology published between 2007 and 2017.

Data sources, eligibility criteria: We undertook searches for peer-reviewed articles in PubMed, Web of Science, Scopus, GreenFile, ScienceDirect and Toxline and included studies of controlled exposures of environmental chemical pollutants, defined as unintentional exposures leading to unintended effects. Studies with viruses, prions or therapeutic agents were excluded, as were records with missing details on chemicals' identities, toxicities, doses, or concentrations.

Study appraisal and synthesis methods: To examine the internal validity of studies we developed a risk-of-bias tool tailored to mixture toxicology. For a subset of 388 entries that claimed synergisms or antagonisms, we conducted a quantitative reappraisal of authors' evaluations by deriving ratios of predicted and observed effective mixture doses (concentrations).

Results: Our searches produced an inventory of 1220 mixture experiments which we subjected to subgroup analyses. Approximately two thirds of studies did not incorporate more than 2 components. Most experiments relied on low-cost assays with readily quantifiable endpoints. Important toxicity outcomes of relevance for human risk assessment (e.g. carcinogenicity, genotoxicity, reproductive toxicity, immunotoxicity, neurotoxicity) were rarely addressed. The proportion of studies that declared additivity, synergism or antagonisms was approximately equal (one quarter each); the remaining quarter arrived at different evaluations. About half of the 1220 entries were rated as "definitely" or "probably" low risk of bias. Strikingly, relatively few claims of synergistic or antagonistic effects stood up to scrutiny in terms of deviations from expected additivity that exceed the boundaries of acceptable between-study variability. In most cases, the observed mixture doses were not more than two-fold higher or lower than the predicted additive doses. Twenty percent of the entries (N = 78) reported synergisms in excess of that degree of deviation. Our efforts of pinpointing specific factors that predispose to synergistic interactions confirmed previous concerns about the synergistic potential of combinations of triazine, azole and pyrethroid pesticides at environmentally relevant doses. New evidence of synergisms with endocrine

Abbreviations: CA, Concentration Addition; DA, Dose Addition; IA, Independent Action; IPQ, Index of Prediction Quality; PECO, Population, Exposure, Comparator, Outcome; PRISMA-P, Preferred Reporting Items for Systematic review and Meta-Analysis Protocols; RoB, Risk of Bias; TUS, Toxic Unit Summation.

* Corresponding author at: Brunel University London, Institute of Environment, Health and Societies, Kingston Lane, Uxbridge, Middlesex UB8 3PH, United Kingdom.

E-mail address: andreas.kortenkamp@brunel.ac.uk (A. Kortenkamp).

¹ These authors contributed equally to the work.

<https://doi.org/10.1016/j.envint.2020.106206>

Received 19 June 2020; Received in revised form 8 October 2020; Accepted 9 October 2020

Available online 26 October 2020

0160-4120/© 2020 The Authors. Published by Elsevier Ltd. This is an open access article under the CC BY license (<http://creativecommons.org/licenses/by/4.0/>).

disrupting chemicals and metal compounds such as chromium (VI) and nickel in combination with cadmium has emerged.

Conclusions, limitations and implications: These specific cases of synergisms apart, our results confirm the utility of default application of the dose (concentration) addition concept for predictive assessments of simultaneous exposures to multiple chemicals. However, this strategy must be complemented by an awareness of the synergistic potential of specific classes of chemicals. Our conclusions only apply to the chemical space captured in published mixture studies which is biased towards relatively well-researched chemicals.

Systematic review registration number: The final protocol was published on the open-access repository Zenodo and attributed the following digital object identifier, doi: <https://doi.org/10.5281/zenodo.1319759> (<https://zenodo.org/record/1319759#.XXIzdy7dsqM>).

1. Introduction

Several previous reviews of synergisms and antagonisms in chemical mixtures (Belden et al. 2007; Boobis et al. 2011; Cedergreen 2014; Deneer 2000; Grenier and Oswald 2011; Vijver et al. 2011; Warne and Hawker 1995) have assessed the utility of concepts for the prediction of mixture effects based on the toxicity of individual mixture components. These efforts focused on specific chemical mixtures and toxicity endpoints. Warne and Hawker (1995) reviewed mixtures of chemicals with an unspecific, narcotic mode of action on aquatic organisms; other toxicities were not considered. Deneer (2000) limited himself to mixtures of pesticides and their effects on aquatic organisms. Belden et al. (2007) reviewed papers on mixtures of pesticides which detailed 303 separate mixture experiments. Most of the studies were for ecotoxicological endpoints; mixtures of other chemicals were not included. Vijver et al. (2011) looked only at mixtures of Cd, Cu or Zn and focused on ecotoxicological studies with water-exposed organisms. Grenier and Oswald (2011) analysed 112 records describing exposure of laboratory or farm animals to combinations of mycotoxins. Their review only considered binary mixtures in a 2×2 factorial design. Cedergreen (2014) is a review of mixture effects with ecotoxicological endpoints. It is an amalgamation of the Belden et al. (2007) and the Vijver et al. (2011) datasets, with additional searches on mixtures of antifoulants and an update with papers published up to 2013. Altogether, 351 papers were analysed. Boobis et al. (2011) focused exclusively on the question of synergisms at low doses in studies of relevance to human/mammalian toxicology.

These reviews concluded that synergisms or antagonisms are relatively rare. Warne and Hawker (1995) did not observe deviations that exceeded predicted effect concentrations by more than a factor of 3. Deneer (2000) and Belden et al. (2007) found that around 90% of experiments that evaluated dose additivity observed effect doses within a factor of 2 of predicted (additive) values. Similarly, in Cedergreen's review (2014), synergistic interactions occurred in a minority of cases, but relatively frequently (26%) with mixtures of antifoulants. The exception is Vijver et al. (2011) where the dominant pattern was one of interactions, with antagonisms the most frequent. Additivity rarely occurred. Boobis et al. (2011) found that the magnitude of synergisms at low doses did not exceed effect doses predicted for additivity by more than a factor of 4.

While the trends observed in these reviews are informative, doubts remain as to whether they can be generalised. Missing is a systematic review of mixtures relevant to human/mammalian toxicity endpoints. Furthermore, a review of all groups of chemicals included in mixture experiments is required, beyond pesticides, metals, mycotoxins or antifoulants. A systematic analysis of risk-of-bias during the formulation of mixture effect predictions, their experimental testing and the final assessment is also missing, as is a systematic quantitative reappraisal of study author claims of deviations from expected additivity.

Our review fills these gaps. We provide material that can support a better-grounded assessment of the extent and frequency of synergisms or antagonisms for a wider range of toxicity endpoints, applicable to all environmental chemicals. This is timely as the current practice of assessing individual chemicals in isolation, without considering risks

associated with combined exposures, is increasingly questioned for being insufficiently protective (Drakvik et al., 2020; Kortenkamp and Faust, 2018; Evans et al., 2016). If mixture effects can be approximated by using prediction tools that assume additivity, the assessment of combined exposures in regulatory practice becomes feasible. However, if synergisms occur frequently, such tools have limited utility and would have to be replaced with specific assessment strategies and regulatory approaches aimed at safeguarding against synergisms more generally (Bopp et al. 2015; European Commission (EC), 2012; Kienzler et al. 2014).

2. Methods

2.1. Definition of key terms

Synergy, synergism, antagonism or synergistic/antagonistic toxicological interactions denote deviations from mixture effects predicted under the assumption that mixture components produce toxicity without interacting with each other, the non-interaction or additivity assumption. Where mixture components are active alone, the non-interaction assumption must be specified in terms of an appropriate additivity model, such as Dose Addition (DA) (also called Concentration Addition, CA), Independent Action (IA) or a mixed model. In this situation, synergy has the specific meaning of "more-than-additive" and antagonism of "less-than-additive".

Where only one component in a binary mixture is effective, non-interaction means that the mixture is not more toxic than that active component alone. In this situation, synergy is synonymous with potentiation and has the specific meaning of an increased effect caused by the second agent, while a lack of influence is usually denoted as "inertism". An analogous definition applies to antagonism. The terms *potentiation*, *effect enhancement*, *toxic enhancement* or *positive effect modulation* are reserved for combinations where one or several chemicals exacerbate the toxic effects of other substances, but without producing that effect on its own. Examples would be the toxicity enhancements observed with piperonyl butoxide and pyrethroids in insects (Cedergreen 2014). In the case of binary combinations, where only one component is active, and the mixture is more toxic than the active chemical, the terms potentiation, effect enhancement, toxic enhancement and positive effect modulation are used synonymously with synergism.

Interaction (mixture interaction or toxic interaction) is used to denote the phenomenon of any deviations from the non-interaction or additivity assumption for combinations of chemicals. This includes deviations from additivity in the case of components that all produce a common toxic effect and deviations from inertism (as defined above) in the case of components that are inactive on their own. Such interactions can be either synergistic or antagonistic, i.e. with effects stronger or weaker than expected under the additivity null hypothesis. The term interaction is purely descriptive and does not imply anything about the nature of the underlying mechanisms.

2.2. Protocol, eligibility criteria and information sources

In this review we identify the extent of deviations from expected

additivity and pinpoint chemical mixtures prone to interactions. Table 1 details the primary and secondary PECO statements and eligibility criteria. Table 2.

The protocol was drafted according to the PRISMA-P (Preferred Reporting Items for Systematic review and Meta-Analysis Protocols)

Table 1
PECO statements and eligibility criteria.

PECO Statement	Inclusion criteria	Exclusion criteria
Populations	All living organisms (covering vertebrates, invertebrates, algae/plants, microorganisms) at all levels of biological organisation (<i>in vivo</i> and <i>in vitro</i> , from cell to mesocosms)	- Viruses, prions
Exposures	Experimental mixtures of all types of man-made chemicals (industrial chemicals, pesticides, biocides, pharmaceuticals, cosmetic ingredients, environmental and food contaminants etc.) of known composition and for which the toxicity of individual component is known/ documented.	Controlled exposures: - known chemical identity, - known concentration, - known toxicity for the combined effect of interest, - at least one compound in the mixture meets the definition for 'environmental chemical pollutant'. - Observational studies, - unknown chemicals, - naturally occurring compounds (of natural origin, in expected place at expected concentration), - unknown concentration, - unknown toxicity for the combination effect under investigation, - combination of chemical and non-chemical stressors (e.g. radiation).
Comparators	Additivity based prediction(s) of combined effect from single substance testing	The above must be true for all mixture components. Information on the chemicals' identity, concentration and toxicity above is missing for one or more mixture component.
Outcomes	<u>Primary outcomes</u> - Concentration(s) of individual components, - their chemical class and/or structure, - their use(s), or regulatory regime under which they fall (if applicable) - environmental media/routes of exposure, - whether observed mixture effects deviate from additivity predictions, <u>Secondary outcome</u> - the direction and magnitude of deviation from additivity based combined effect predictions.	Comparison of a mixture observation with a mixture expectation Intended effects (e. g. therapeutic effects of pharmaceutical or dietary supplement) following intentional exposure

Table 2

Results of the literature search by source.

Search string ¹	Source	Results
(MixTox1 OR Characterisation1) OR (Mixtox2 AND Characterisation2)	PubMed	5593
(MixTox1 OR Characterisation1) OR (Mixtox2 AND Characterisation2)	Web of Science	2977
(MixTox1 OR Characterisation1) OR (Mixtox2 AND Characterisation2)	Scopus	2398
(MixTox1 OR Characterisation1) OR (Mixtox2 AND Characterisation2)	Green file - EBSCO	752
chemical mixture (MixTox1 OR Characterisation1) OR (Mixtox2 AND Characterisation2)	Science direct	2081
Manual searches	DART Toxline	1
Manual searches through reference list of eligible studies	Grey literature	976
	Eligible studies	69

¹ Further details about the search strings used can be found in doi: <https://doi.org/10.5281/zenodo.1319759> (<https://zenodo.org/record/1319759#.XXIzdy7dsqM>).

2015 checklist (Shamseer et al. 2015). We gave due consideration to all eight key elements of the Code of Practice for the Conduct of Systematic Reviews in Toxicology and Environmental Health Research (COSTER) (Whaley et al. 2020). The final protocol was published on the open-access repository Zenodo and attributed the following digital object identifier, doi: <https://doi.org/10.5281/zenodo.1319759> (Martin et al. 2018) (<https://zenodo.org/record/1319759#.XXIzdy7dsqM>).

Boundaries around the chemical space of interest were set by defining an *environmental chemical pollutant* as a substance released into the environment (atmosphere, lithosphere, hydrosphere and biosphere) as a result of anthropogenic activities and found therein in unexpected places and/or in unexpected quantities. This includes mycotoxins in crops but excludes other natural poisons when found in expected places in expected quantities. It also excludes endogenous chemicals produced as a response to environmental stressors.

Environmental exposure was defined as *unintentional exposure* leading to *unintended effect(s)*. Accordingly, drug administration where intended therapeutic effects are of interest, or intentional poisoning if self-harm was the intended effect, were excluded, while environmental exposure to pharmaceuticals was included. Food additives and cosmetic ingredients were included as they are added intentionally by manufacturers to support specific food product properties but may lead to unintended health effects. By default, only studies concerned with the intended effect of intentional use were excluded.

The PECO statements and above definitions were operationalised as inclusion and exclusion criteria, as described in Table 1.

We limited the publication period of interest to literature published between the 1st January 2007 and 1st May 2018 in any language. For languages other than English, French or German, for which no version of the full text in English could be located, we attempted using the online translating tool Google Translate before we excluded an article as 'article not accessible'.

We conducted searches for peer-reviewed articles in the following bibliographic databases:

- PubMed
- Web of Science Core Collection
- Scopus
- GreenFile
- The full text database ScienceDirect
- Toxline + DART (TOXNET platform)

In addition, the CREST database (Chemical attributes, Regulatory approaches and Experimental Studies from a mixture toxicology perspective) built as part of a previous project contains 260 peer-reviewed articles published between 2007 and November 2014 (www.crestdb.org).

rmonline.net/CREST).

In order to identify ongoing research, we examined conference papers via resources like the British Library service **Zetoc** (<http://zetoc.jisc.ac.uk/>).

To identify grey literature, searches were carried out using the topic focused search engines **Environar** (<https://environar.com/environar/desktop/en/search.html>), in open access bibliographical databases such as **OpenGrey** (<http://www.opengrey.eu/>) or **CORE** (<http://www.core.ac.uk>) to search open access items in institutional repositories. This was complemented by targeted manual searches of open repositories on the website of European, American and national institutions as well as those of interest groups (full details in published protocol (Martin et al. 2018)).

2.3. Literature search

Topical vocabulary in the literature describing toxicological assessments of chemical mixtures contains expressions and phrases which have little discriminatory power, such as “combination”, “mixture” or “joint”. This was a major challenge in identifying relevant literature. To devise a specific search strategy that remained sensitive enough to avoid missing important literature, we piloted two lists of terms in Web of Science, PubMed and Scopus; the first list consisted of 17 terms or expressions relating to mixture toxicology generally, the second is a list of 48 terms or expressions related to mixture effects (See (Martin et al. 2018) (<https://zenodo.org/record/1319759#.XXIzdy7dsqM>)). Some terms or expressions were extremely frequent, yielding hundreds of thousands of hits, whilst other were more specific (under 300 hits). The more frequent terms required an additional filter. This was achieved by combining each list with each other using an ‘AND’ Boolean operator, i. e. the title, abstract and key words of a study needs to contain at least one expression related to mixture toxicology generally and another to mixture effect characterization.

2.4. Study selection

The systematic review process was managed with the support of the free online tool CADIMA (<https://www.cadima.info/index.php/area/evidenceSynthesisDatabase>).

Eligibility criteria (Table 1) were applied to the merged reference list by two team members working independently, and in two stages; title and abstract screening and full text screening. A consistency check on the basis of a subsample of 200 references calculated the inter-reviewer agreement as 0.43 on the basis of agreement for all inclusion/exclusion criteria and was considered ‘fair’ based on the measuring agreement of Cochrane (<http://handbook.cochrane.org/> Part 2, chapter 7.2.6). Inconsistencies were reviewed and interpretation of criteria clarified. The reason for exclusion of studies after assessment of the full text was recorded.

Due to the large number of studies considered eligible after this first full text screening step, we considered the influence of stricter eligibility criteria, specifically with respect to the availability of information on the purity of the chemical compounds and the simplicity of the mixture design. However, the purity of compounds was inconsistently reported and the exclusion of studies on that basis was judged inappropriate (e.g. only the supplier is commonly reported). Studies using commercial formulations and those investigating the effects of particulates where particle size was not characterised in the mixture were excluded. Simple mixture design, where individual compounds had only been tested at one concentration or dose, and for which no additivity expectation can be formulated, were also excluded at this second stage of full text screening. We excluded a total of 490 studies by these stricter eligibility criteria.

Multiple reports of the same research (e.g. multiple publications, conference abstracts etc.) were collated as part of the data extraction process as one unit of evidence.

A further 46 eligible studies that had been overlooked through database searches were identified by scanning the references of eligible studies.

The totality of eligible records constitutes the *data extraction inventory* (Supplementary Material S1).

2.5. Data collection and data items

Chemical mixtures vary in terms of composition and mixture ratio and accordingly can be tested by using a variety of experimental designs. Similarly, there are several approaches to judging deviations between predicted and observed mixture response.

To accommodate this diversity, we had to adopt a versatile data extraction process and proceeded in two stages. A ‘front-end’ data extraction Excel template was used to extract data as reported by authors and relevant for all eligible studies (referred to as the *data extraction inventory*) for the purposes of the primary PECO statement. No data interpretation took place at this first stage. The work in this first stage of the data extraction process was distributed between four project team members.

In a second stage, the information extracted was used to identify mixture studies of potential relevance for a quantitative reappraisal in line with the secondary PECO statement. All potential candidates for a quantitative reappraisal were entered into the *reappraisal database* organised as a Microsoft Access database by one team member (MS) (Supplementary Material S2, S3). The database architecture consists of interlinked tables that capture the various aspects of mixture study designs and assessments and data about the individual compounds.

The data extraction Microsoft Excel template for the *data extraction inventory* (S1) was developed iteratively through parallel piloting of key papers by team members. The resulting spreadsheet and accompanying guidance document can be found in the [supplementary information](#). Briefly, it contains the following items:

- *meta-data* (authors, date, journal name or report number, title, abstract, funding)
- information about the study related to
 - o the mixture characteristics including the reported rationale for the selection of compounds in the mixture (such as class of chemical or chemical structures, uses, regulatory regime or exposure route), the mixture size
 - o type of study (*in vivo*, *in vitro*)
 - o the test system (test species or cell type)
 - o the timing for the generation of single substances data
 - o the study design (point design, fixed ratio or ray mixture, surface mixture design etc.)
 - o the way in which the experimental outcomes were reported, together with the authors’ conclusions

Many publications reported several mixture experiments. We therefore recorded as one entry in the data extraction template all combinations of the same set of chemicals in the same study. When several toxicity endpoints had been evaluated, one endpoint, judged to be the apical endpoint, was recorded in the data extraction template and the others were ignored.

2.6. Risk-of-bias assessment of individual studies

Flaws in the design, conduct, analysis, and reporting of experimental mixture studies can lead to erroneous conclusions about toxicological interactions. Toxicological interactions are identified by comparison between an expected (additive, non-interactive) mixture response calculated based on the toxicity of its components and an empirically observed mixture response. The tacit assumption made during such comparisons is that the toxicological evaluations of single mixture components and the mixture are comparable. If this is not the case, the

comparison is biased. All three elements of the assessment process – calculation of the expected mixture effect, experimental observation of mixture effect, and comparison between prediction and observation – can be biased, leading to erroneous claims of interactions. In the worst-case, a real existing interaction might be overlooked.

In the field of evidence-based toxicology, critical appraisal tools for the assessment of the internal validity of toxicological studies of single substances have come into widespread use. These risk-of-bias tools assess studies in terms of e.g. selection bias (random allocation to dose groups, concealment), performance bias (blinding, identical experimental conditions), detection bias (reliability of exposure characterisation or outcome measurements) and more. However, such tools are out of scope for bias considerations when it comes to assessing bias in declaring toxicological interactions in mixtures.

To evaluate the internal validity of mixture studies we therefore developed a tailor-made risk-of-bias tool which addresses the three domains of mixture effect evaluations (mixture effect prediction, experimental testing and comparison of test results and prediction). The tool consists of a system of risk-of-bias questions which require the reviewer to choose between low and high risk-of-bias options on a 4-point scale (Koustas et al. 2013), ranging from “Definitely low risk of bias”, “Probably low risk of bias”, “Probably high risk of bias” to “Definitely high risk of bias”. An overall risk-of-bias scale for each study was derived by assigning the lowest scale achieved in any one of the three domains. The tool was written in Microsoft Excel and can be found together with an accompanying guidance document on the open-access repository Zenodo (Martin et al. 2018). For convenience, it is also included in the [Supplementary Material](#) to this paper (files S4 and S5). For the avoidance of misunderstandings, we emphasise that our tool does not assess aspects of internal validity already covered by available risk-of-bias tools for single chemical studies.

2.7. Summary measures: Quantifications of toxicological interactions and decision rules for identifying synergisms and antagonisms

A deviation between a predicted and observed mixture effect can be evaluated in several ways: one approach defines the magnitude of an interaction as the ratio of a predicted and observed response at a specified dose. This method evaluates how far the observed dose–response relationship shifts up or down along the effect scale relative to the curve corresponding to expected additivity. In the second approach, the magnitude of a deviation from expected additivity is expressed as the ratio of a predicted and observed dose that cause a specified response (effect dose). With the latter method, shifts of the observed mixture dose–response curves from the expected curve along the dose axis are evaluated.

Comparisons along the response scale, according to the first approach, are fraught with difficulties. This method places great demands on controlling response variations and requires highly reproducible experimental setups. Response variations become large with steep dose–response relationships where small changes in the dose provoke large effect changes. This complicates the comparison of repeated experimental studies, especially when mixture experiments rely on historical data for single substance testing generated a long time before the mixture experiment. In contrast, changes in terms of predicted and observed effect doses, along the dose axis, are usually less pronounced and their experimental evaluation is more robust. For this reason, we chose the latter approach and identified deviations from expected (additive) mixture effects by comparing predicted and observed effect doses (concentrations):

$$\text{Ratio} = \frac{\text{observed } EC_X(\text{mixture})}{\text{expected } EC_X(\text{mixture})} \quad (1)$$

where *observed* $EC_X(\text{mixture})$ is the effective concentration of the mixture obtained from the toxicity experiment and *expected*

$EC_X(\text{mixture})$ the concentration of the mixture that was predicted by the additivity model to produce an effect X . This method expresses an observed mixture effect concentration as a fold difference to the prediction: a value of 2 means that the observed effect concentration is twice the predicted one; and a value of 0.5 means that the observed effect concentration is half the predicted one. Thus, the ratio assigns synergistic mixtures values from 0 to < 1 and antagonistic mixtures values from greater than 1 to infinity.

If CA is used as a reference model for additivity, the ratio in equation (1) is equivalent to the so-called Toxic Unit Summation (TUS):

$$\text{TUS} = \frac{\text{observed } EC_X(\text{mixture})}{\text{expected } EC_X(\text{mixture})} = \sum_{i=1}^n \frac{c_i}{EC_{Xi}} \quad (2)$$

where c_i is the concentration of the i^{th} compound in the n -compound mixture that has produced a mixture response X , and EC_{Xi} the concentration of the i^{th} compound leading to the same response X as observed for the mixture.

The ratios defined in equations (1) and (2) are asymmetrical; ratios of e.g. 0.8 and 1.2 do not express the same extent of under- or overestimation. To compensate for these asymmetries, we used the Index of Prediction Quality (IPQ) (equivalent to the Additivity Index by (Marking and Mauck (1975) which operates on a linearized symmetric scale and expresses an over- or underestimation as a percentage of the ratio in equation (2):

$$\text{IPQ} = \begin{cases} 100 * (\text{Ratio} - 1) & \text{if } \text{Ratio} \geq 1 \\ 100 * \left(1 - \frac{1}{\text{Ratio}}\right) & \text{if } \text{Ratio} < 1 \end{cases} \quad (3)$$

An IPQ = 50% means that the observed effect concentration of the mixture is 50% above the predicted value, and conversely, an IPQ = -50% refers to an observed effect concentration of the mixture that is 50% below the predicted value.

Although the IPQ quantifies deviations from predicted additivity, it does not provide criteria for deciding which degree of deviation should be classed as synergism or antagonism. Small deviations may not represent real existing toxicological interactions as they could be the result of experimental error. The issue can be addressed by statistical significance testing. Detection of statistically significant deviations from predicted additivity can guide assessments of mixture effects in terms of synergisms (negative IPQ values) or antagonisms (positive IPQ values). With a study of good data quality, even small differences can become statistically significant. However, it is also important to consider between-study variations, and a disregard of this source of variation may lead to underestimations of the true uncertainty assessments in terms of synergisms or antagonisms. Thus, exclusive reliance on statistical significance tests in deciding on synergisms or antagonisms may be misleading.

For these reasons, pragmatic decision rules that take account of both statistical and experimental issues have been applied in previous reviews. According to Belden et al. (2007), all IPQs that fall into a range between -100% and 100% are assessed as additive. Broderius et al. (1995) have used a narrower classification and consider an IPQ as additive only if it falls within the range of -25% to 20% (corresponding to a TUS range of 0.8 to 1.2). In our quantitative assessments, we have used both Belden et al. (2007) and Broderius et al. (1995).

2.8. Quantitative reappraisal of reported toxicological interactions

All records in the *data extraction inventory* (see [Supplementary Material S1](#)) which showed evidence for interactions in relation to DA (CA) expectations were selected for a quantitative reappraisal. The selected records constituted our *reappraisal database* ([Supplementary Material S2](#)), together with a user manual, [Supplementary Material S3](#)). All studies in the *data extraction inventory* which were assessed as additive in relation to DA (CA) by their authors and where our risk-of-bias assessment

signalled “definitely” or “probably” low risk were excluded from the *reappraisal database*. We also excluded studies where a quantitative appraisal of interactions was not possible due to inappropriate experimental design or gaps in the reported data. Further, we disregarded studies where the authors did not assess data with respect to an additivity hypothesis but where agreement with DA (CA) predictions was evident (for further details about the methods used in our quantitative reappraisal see [Supplementary Material S6](#)). Cases where authors did not draw any conclusions about additivity but where the data suggested interactions, were included in the *reappraisal database*.

A quantitative reappraisal of mixture studies based on effect doses (concentrations) as required for calculating IPQs is difficult when IA is used as the assessment concept. Studies employing IA often evaluate deviations from expected mixture effects in terms of shifts along the

effect axis. Not only are effect doses for IA rarely reported in mixture studies, but recalculations of IA mixture predictions require that complete dose–response descriptions be available for every component in the mixture. Furthermore, when the number of components in the mixture increases, mixture effect predictions derived from IA must rely on high quality dose–response data in the low dose range. Most studies in the *data extraction inventory* did not meet these data requirements. For this reason, we did not attempt reappraisals of mixture studies based on IA. If, however, a mixture expectation according to IA was reported as an effect dose (concentration), we recorded the deviation by calculating TUS or IPQ according to Equations (2) and (3).

Decisions to select a mixture experiment for inclusion in the reappraisal database were driven by the data requirements that had to be met to permit quantitative reappraisals. Based on these data requirements

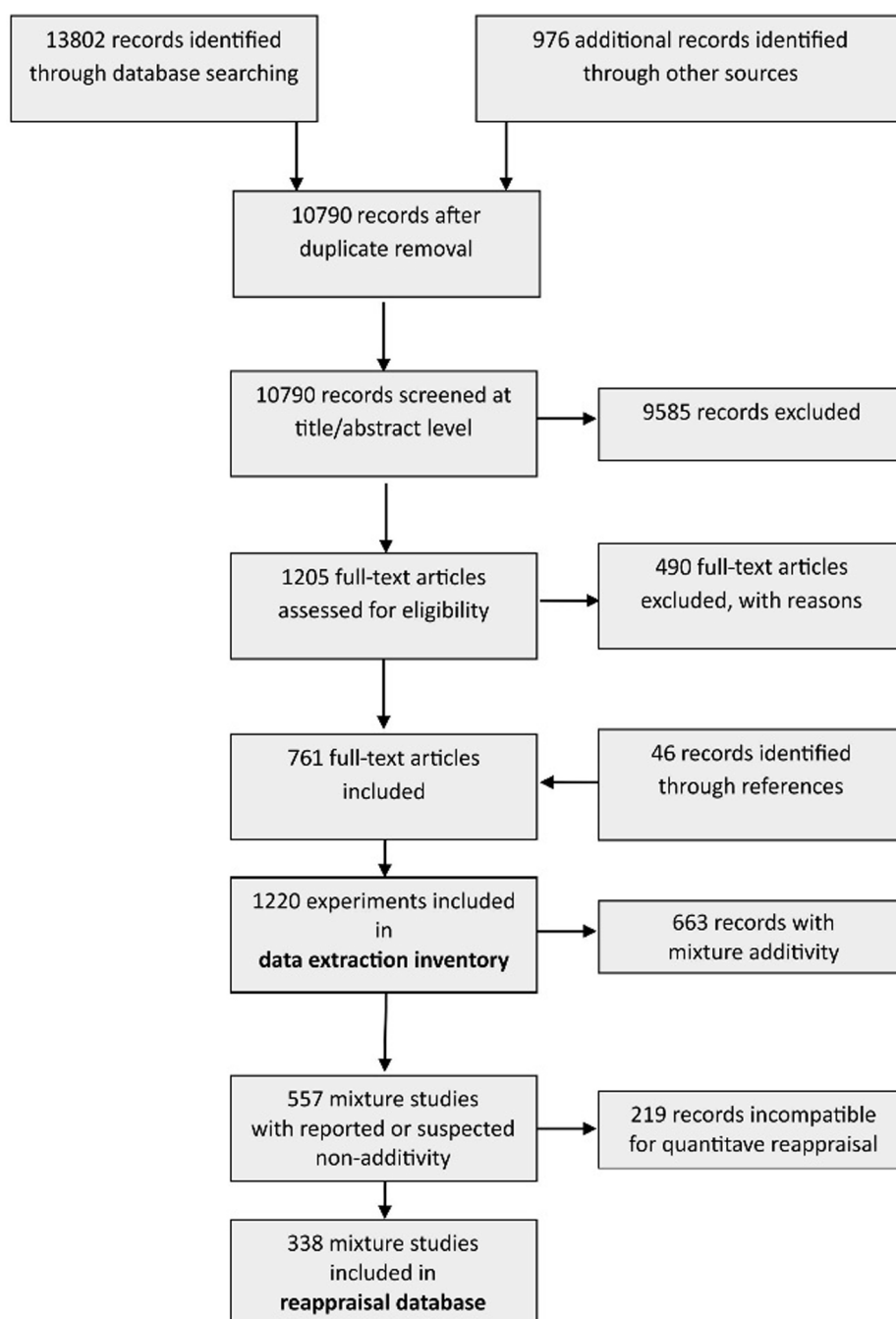


Fig. 1. Flow diagram of study selection for entry into the inventory of mixture studies.

we derived criteria for inclusion. Both, data requirements and the resulting criteria, are described in [Supplementary Material S4](#). Often, outcomes for more than one mixture were reported, or mixture responses for several mixture ratios of the same mixture were available. It was then possible to calculate several TUS for the same effect level and endpoint. In such cases, the TUS with the largest deviation from additivity was chosen, and therefore maximally only one entry per record of the data extraction inventory entered the reappraisal database.

Where the criteria for inclusion were not met, quantitative reappraisals could not be conducted, and we flagged the corresponding mixture entry in the inventory as unsuitable for reappraisal. When a TUS (or IPQ) was reported by study authors its value was taken from the record and directly entered into the database. We re-calculated the TUS where it was not reported or where our risk-of-bias evaluation indicated that its determination might have been done improperly (i.e. when we judged that the comparative assessment of observed and predicted mixture effects was not supported by data evidence and analysis, see the RoB in [Supplementary Material S1](#)).

Due to the focus on CA and the data requirements for a TUS calculation, not all types of synergism could be reappraised. For example, the enhancement of a substance's potency in the presence of an inactive substance (potentiation) was not covered by the appraisal criteria and was marked as non-appraisable.

All association analyses were performed on categorical variables by the Fisher's exact test and the Chi-square with Yates continuity correction.

3. Results

Our literature searches identified 13,802 records in databases and an additional 976 records through manual searches ([Fig. 1](#), [Table 1](#)). After duplicate removal, 10,790 records were screened for relevance by scrutiny of titles and abstracts. This process excluded 9585 records. The full texts of the remaining 1205 records were examined for eligibility for inclusion in the inventory of mixture studies which led to the exclusion of 490 records. During the appraisal of full-text articles, 69 additional records were identified via analysis of the references and 46 of those were subsequently included in the inventory. Finally, data was extracted from 761 records. Because several records described more than one mixture experiment, this is equivalent to 1220 experiments ([Fig. 1](#)).

3.1. Characteristics of eligible studies

The number of studies eligible for entry into the *data extraction inventory* increased after 2010. The most frequently reported outcome was additivity (28.3%), followed by synergism (24.3%) and antagonism (19.2%). In 28.2% of the studies, authors declared interactions, potentiations or other assessments ([Fig. 2](#)). Most studies (87%) were publicly funded, with a minority (1.8%) supported by industry or private sources. In 7.8% of all studies, funding sources were not disclosed. The stated aims of 89% of all studies were "proof of principle", i.e. as assessments of the predictability of mixture effects based on the toxicity of all mixture components, in accordance with our eligibility criteria. Only 6.3%

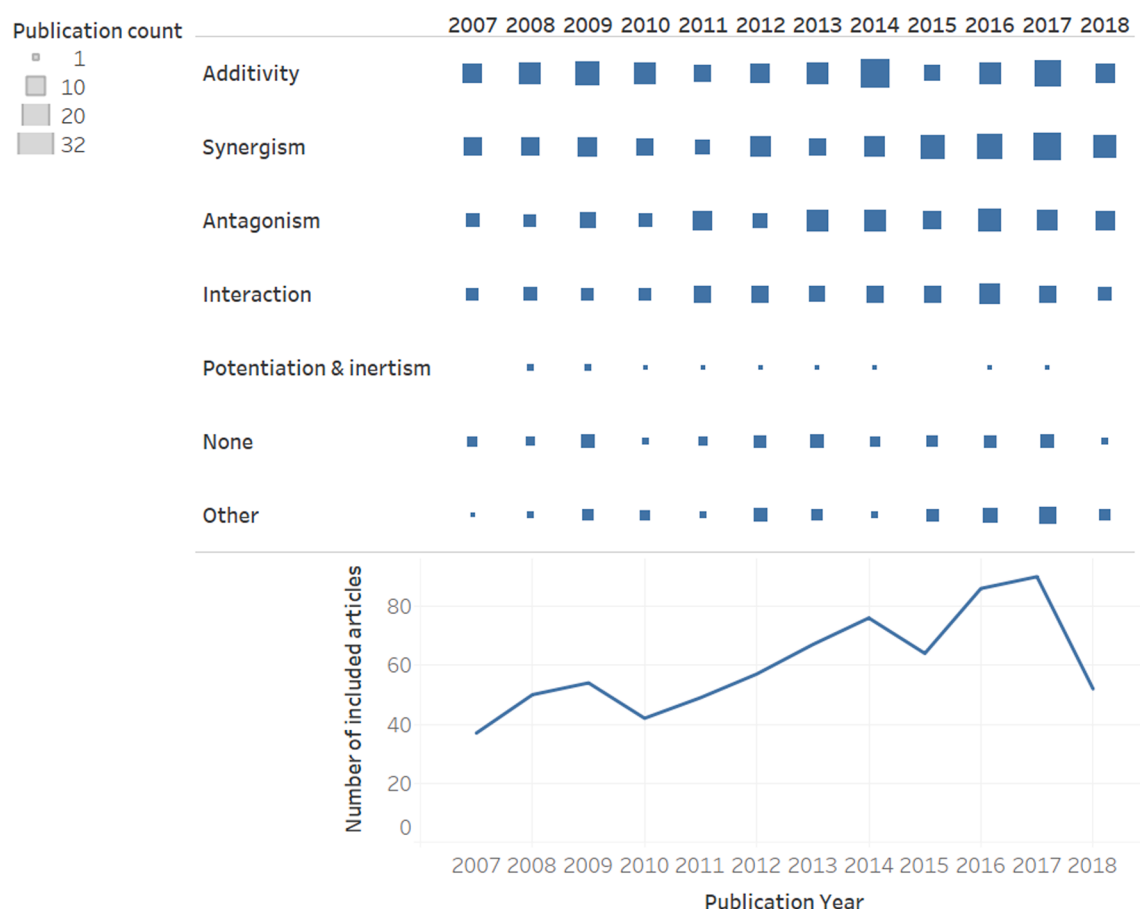


Fig. 2. Number of studies eligible for entry into the inventory by year and reported study outcome (until 1st May 2018). "None" refers to experiments for which authors did not assess mixture effects and "others" included a diverse category where authors had commented on mixture effects using concepts not directly related to additivity (e.g. microbial agonist/antagonist activity) (Interactive view).

https://public.tableau.com/views/JRCMixture/Public-yearclaim?:language=en-GB&:display_count=y&:origin=viz_share_link

examined mixture effects at environmentally relevant levels, while a vanishingly small proportion evaluated mixture toxicity when all components were combined at low doses. More than 99% of all studies investigated the effects of simultaneous exposure to multiple chemicals. Only 0.25% evaluated joint toxicity after sequential exposure.

Most eligible mixture studies employed low-cost assays with readily quantifiable endpoints (Fig. 3), such as acute systemic toxicity and mortality/lethality in ecotoxicological mixture studies. In mixture studies from the human/mammalian toxicology domain, assays related to cytotoxicity and endocrine disruption (*in vitro*) dominated. Presumably due to their high cost, mixture studies that investigated endpoints related to carcinogenicity, genotoxicity and mutagenicity were rare, in both ecotoxicology and human/mammalian toxicology. Similarly, immunotoxicological and neurotoxicological studies are vastly under-represented. In human/ mammalian toxicology, studies with *in vitro* assays predominated. This was not the case in ecotoxicology, due to different conventions in categorising assays in terms of *in vitro* and *in vivo*. Investigations in some taxonomic groups, such as amphibians and birds were under-represented. The majority of ecotoxicological mixture studies administered test compounds by exposure through water. The next most frequent mode of delivery was *in vitro*. Among *in vivo* studies, oral administration was the most frequent exposure route. Mixture toxicity after application by other routes, such as by inhalation or through the skin, was only rarely investigated.

About 80% of studies in the *data extraction inventory* incorporated 2 or 3 mixture components, with 62% involving binary mixtures. Mixture experiments with more than 12 components were uncommon (Fig. 4). By far the most frequently used design (53.5%) was the fixed mixture ratio approach. About 7% of studies used a simple design in which several chemicals were tested at fixed doses, followed by combining all substances at the same doses in a mixture (single dose summing up).

Approximately the same proportion of studies employed a set up where one chemical in the combination was held fixed, while the doses of others were escalated (A in the presence of B). Around 12% of experiments adopted the isobologram approach. Approximately 5% of experiments constructed response surfaces where two substances are combined at various mixture ratios and the resulting mixture effects captured as a response surface in 3-dimensional representations. In about 4% of experiments, the adopted design could not be identified, and over 10% of studies used a variety of approaches different from the other designs.

The categories of chemicals investigated most frequently in mixture experiments were active substances in pesticidal and biocidal products, metal compounds and pharmaceuticals. Interactions, and specifically antagonisms appeared to be reported more often for metals.

The proportion of additivity assessments was largest among endocrine disruptor studies and experiments investigating cell proliferation (42% and 53%, respectively, Fig. 5). Seventy-five percent of carcinogenicity mixture experiments returned assessments of synergy; however absolute numbers were small. The number of chemicals incorporated in mixtures had an influence on the mixture effect as reported by study authors. The larger the number of components incorporated in mixtures, the larger the fraction of studies reporting additivity (Fig. 4). However, the number of such multi-component mixtures was too small to usefully comment on the validity of the funnel hypothesis (Warne and Hawker 1995) beyond the fact that the evidence available from our *data extraction inventory* does not disprove it. The experimental design also had an influence on the reported mixture effect. The proportion of studies declaring deviations from additivity was lowest for the fixed mixture ratio design (62%) and highest for “single dose summing up” (92%).

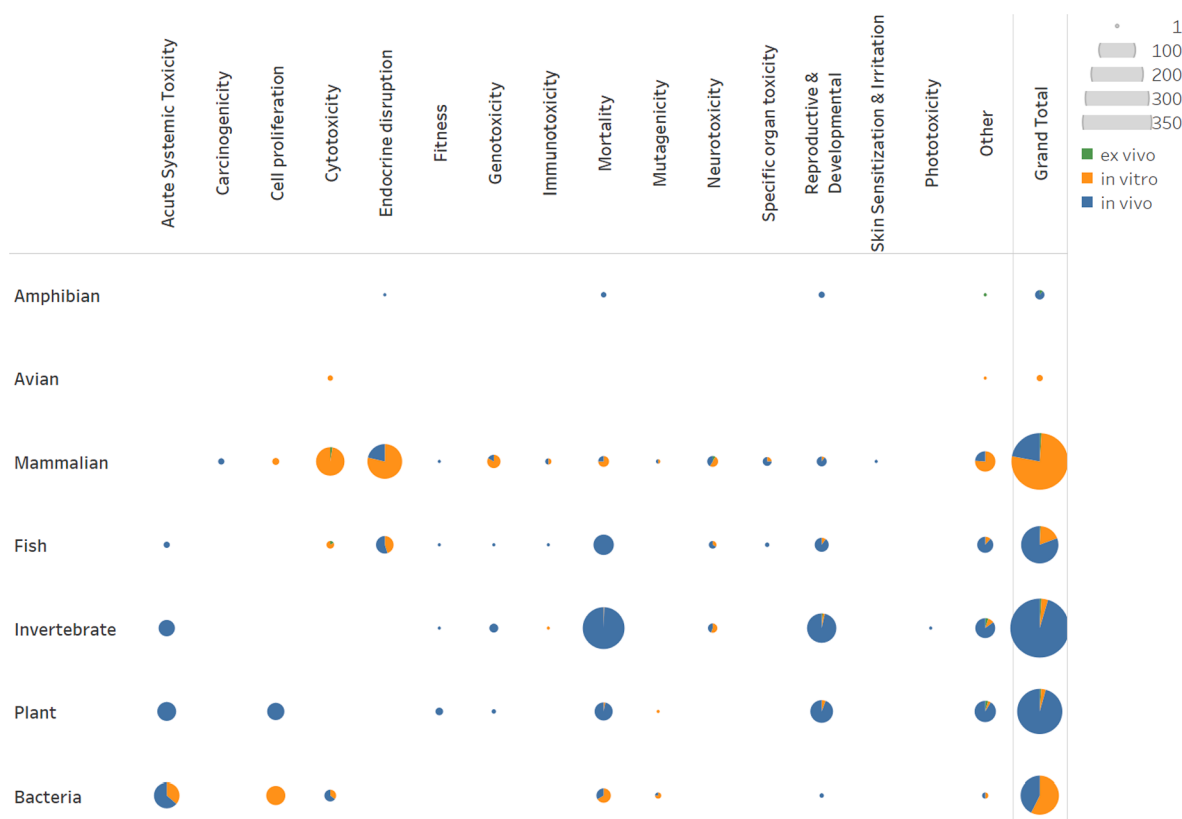


Fig. 3. Number of mixture studies by taxonomic group and endpoint (Interactive view).

https://public.tableau.com/views/JRCMixture/Public-TaxonEndpoint?:language=en-GB&:display_count=y&:origin=viz_share_link

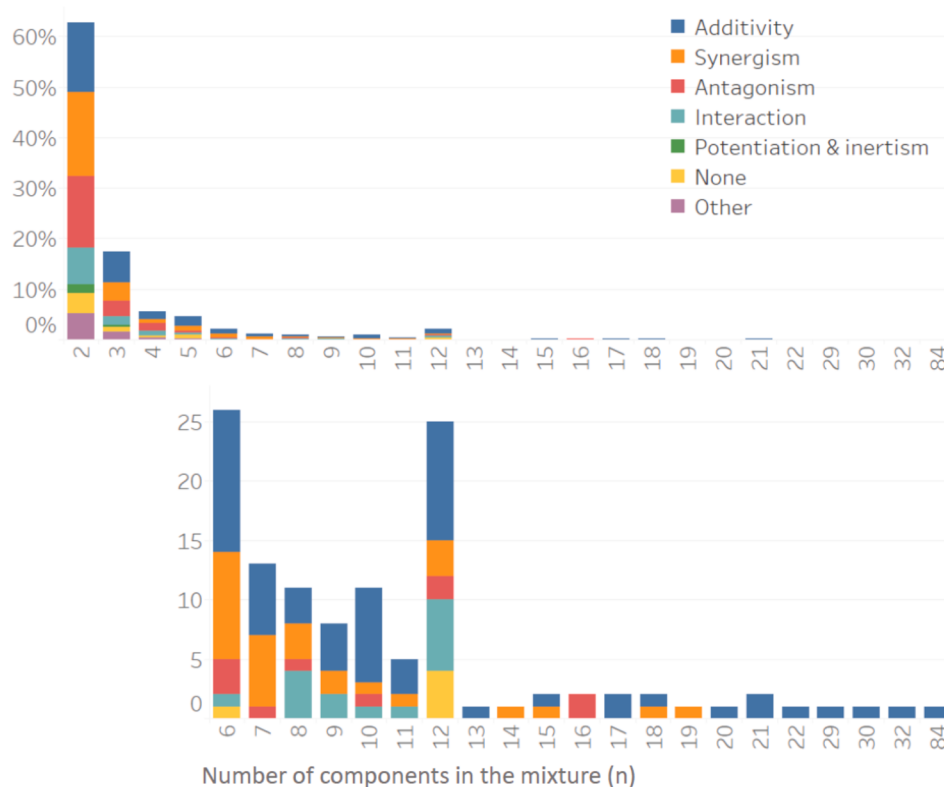


Fig. 4. Reported mixture effects by number of mixture components as a percentage of the total number of mixture experiments (top), and as the number of mixture experiments for mixtures with more than 5 components (bottom) (Interactive view).

https://public.tableau.com/views/JRCMixture/Public-Mixturesizeandassessment?:language=en-GB&:display_count=y&:origin=viz_share_link

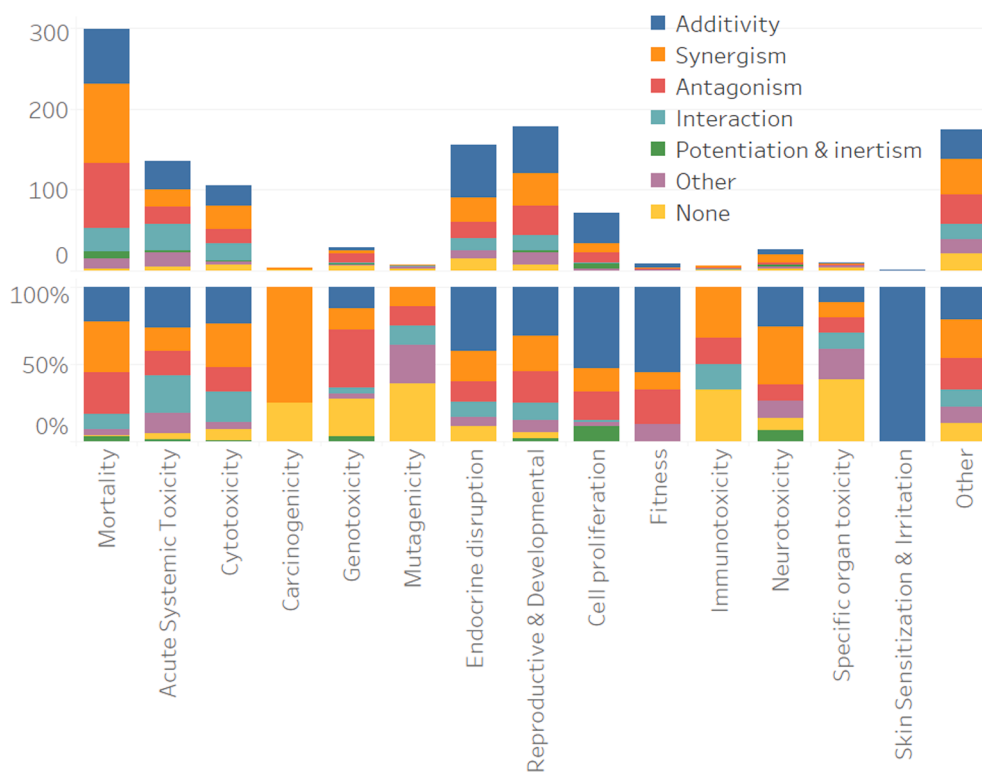


Fig. 5. Mixture studies by toxicity endpoint and reported type of mixture effect by number of mixture experiments (top), and as a proportion of the total number of mixture experiment per endpoint (bottom). (Interactive view).

https://public.tableau.com/views/JRCMixture/Public-Mixtureendpointandassessment?:language=en-GB&:display_count=y&:origin=viz_share_link

3.2. Risk-of-bias within studies

We evaluated the internal consistency of mixture studies in the *data extraction inventory* in terms of three domains: The “mixture expectation” domain assesses studies in relation to the way in which additivity expectations were derived. The “mixture observation” domain rates the reliability of experimental measurements of mixture effects, while the “Comparative assessment” domain evaluates the consistency of comparisons between expectation and observation. We rated approximately 50% of studies as “probably low risk” of bias, about 30% as “probably high risk”, 18% as “definitely high risk” and the remainder as “definitely low risk” (Fig. 6). These assessments were driven by shortcomings in the “mixture expectation” and “mixture observation” domains, and less so in the “comparative assessment” domain. The “probably high risk” assessment category was driven by shortcomings in all domains (Supplementary material). The judgement “definitely high risk” was strongly influenced by deficiencies in the “comparative assessment” domain, but less so in the other domains. There were no clear differences related to study endpoints and nature of effect, or whether mixture studies were ecotoxicological or from the human/mammalian toxicology.

In approximately 40% of the experiments rated as “definitely low risk”, mixture effects had been evaluated by authors as additive. Among the studies assessed as “definitely high risk”, this proportion decreased to around 6%. Furthermore, the “definitely high risk” class had the

largest proportion of unclear or ambiguous mixture effect evaluations (“other” or “none”) (Fig. 6).

3.3. The quantitative reappraisal database and its characteristics

Of the 1220 entries in the *data extraction inventory*, 557 claimed deviations from expected additivity and were classed by their authors as synergisms, antagonism, interactions or potentiations. We considered these 557 entries as candidates for a quantitative reappraisal (Fig. 1). For ca. 70% (N = 388) of these records it was possible to re-calculate mixture expectations according to DA (CA) and to derive IPQs. Together, these 388 entries constitute what we refer to as the *reappraisal database* (see Fig. 1).

However, for 169 of the 557 candidates from the *data extraction inventory* a quantitative reappraisal was not possible, for one or several of the following reasons: (i) the mixture composition was not reported, (ii) data for responses for the same effect magnitude were missing, both for individual mixture components and the mixture itself, (iii) data were recorded only as graphs of poor resolution which made data readouts impossible, (iv) inadequate mixture designs were used (e.g. factorial designs with missing dose allocations), and (vi) inadequately described outputs from software tools were recorded.

We identified 34 mixtures where their authors utilised IA as the additivity assessment concept. In 6 of these 34 cases, the recorded

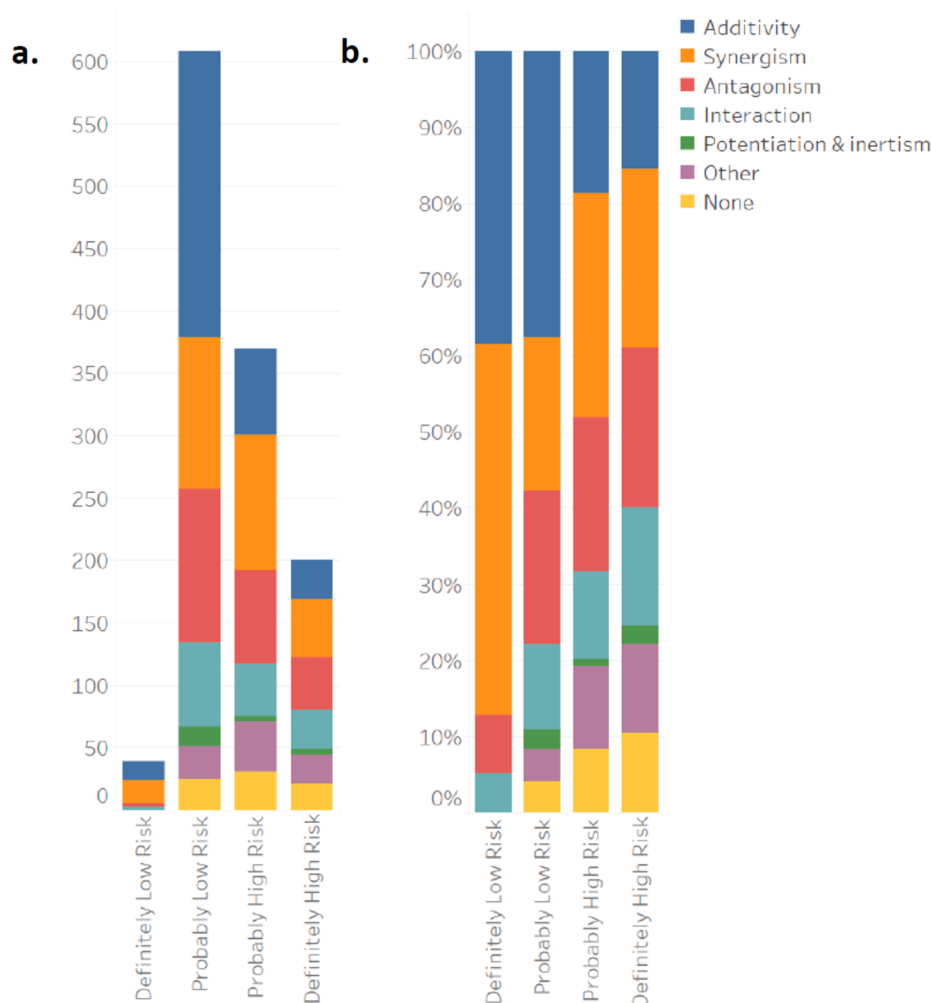


Fig. 6. Authors' claims of type of mixture effect as influenced by risk-of-bias as, a) absolute counts or, b) proportions. (Interactive view). https://public.tableau.com/views/JRCMixture/Public-RiskofBias?:language=en-GB&:display_count=y&:origin=viz_share_link

Table 3

Overall risk-of-bias scores among studies excluded or included in the reappraisal database.

	Definitely High Risk	Probably High Risk	Probably Low Risk	Definitely Low Risk	Total
Included	30 (7.7%)	132 (34.0%)	217 (56.0%)	9 (2.3%)	388 (100%)
Excluded	57 (33.6%)	54 (32.0%)	54 (32.0%)	4 (2.4%)	169 (100%)
Total	87	186	271	13	557

mixture responses fell within the window defined by the predictions derived from CA and IA. In 27 of the cases, CA predicted a higher mixture toxicity than IA. These prediction differences have an impact on the assessment of mixture effects in terms of synergism or additivity. What will be assessed as additive according to CA will be identified as synergistic based on IA. Therefore, with IA and the Belden criteria, we assessed 21 mixtures as synergistic, but only 8 with CA as the reference.

Of the 169 entries excluded from the *reappraisal database*, nearly one third scored as “definitely high risk” in our risk-of-bias assessment. In contrast, only 7.7% of all the 388 records for which IPQs could be established received this score (Table 3).

Twenty three % (N = 88) of the 388 entries in the *reappraisal database* were for mixture experiments relevant to the human/mammalian toxicology domain and 78% (N = 303) were tested in ecotoxicological test systems (for three mixture studies a classification of assays into human or environment was not possible and they were counted in both categories). 100 mixture studies were conducted in *in vitro* test systems, and 285 entries were for *in vivo* bioassays (for three mixture studies a classification in terms of *in vitro* and *in vivo* was not possible). Of the 285 mixture experiments that employed *in vivo* assays, the majority (75%, N = 213) used short-term exposure conditions with acute endpoints. The remainder (N = 72) utilised (sub)chronic tests, the majority of which employed the 72 h growth inhibition test for freshwater alga and cyanobacteria. Only four entries from the domain of human/mammalian studies tested subchronic conditions. Nearly 65% of all entries (N = 249) were experiments involving binary combinations, 17% were ternary and 6% quaternary mixtures. Only a few entries involved more than 4 mixture components (Fig. 4). More than 600 substances were included as a component in at least one mixture. Most mixtures were composed exclusively of pesticides or biocides (22.2%), metals (16.2%) or pharmaceuticals (12.9%). Twenty-three % of mixtures were composed of chemicals that were neither pesticides nor metals (Table 4).

3.4. Quantitative reappraisal of claims of synergisms and antagonisms

The distribution of all 388 IPQ values that we calculated during our reappraisal had a median of −11%, indicating a small underestimation of predicted additive effect concentrations. The 25th percentile was −92% and the 75th percentile 70%. The extremes are IPQ values of −9900% and 838%. The distribution of IPQs was symmetrical which implies that the number of mixture experiments with underestimations of predicted additive effect doses was nearly the same as the number producing overestimations. Division of the IPQs into nine classes centred around IPQ = 0% illustrates the symmetric shape of the IPQ distribution (Fig. 7).

In evaluating these IPQ values in terms of synergisms and

Table 4

Composition of mixtures in the reappraisal database by major chemical types.

	Only metals	Only pesticides or biocides	Only Endocrine disruptors	Only pharmaceuticals	Only mycotoxins	Neither pesticides nor metals
N	62	86	22	50	22	92
%	16.0	22.2	5.7	12.9	5.7	23.7

antagonisms we followed Belden et al. (2007) who class IPQs that fall between −100% and 100% as additive. Using this criterion, we assessed 65% (N = 252) of all reappraised mixture experiments as additive, 20% (N = 78) as synergistic and 15% (N = 58) as antagonistic.

According to the stricter criteria proposed by Broderius et al. (1995), IPQs between −25% and 20% are evaluated as additive. Application of the Broderius criteria leads to the classification of 19% of the 388 entries as additive.

Most mixture studies selected for entry into the *reappraisal database* were evaluated as “probably low” or “probably high” risk of bias (N = 349, Table 5). For around 10% of entries the score was “definitely high” risk of bias. Associations between risk-of-bias scores and classifications in terms of synergisms (IPQ < −100%) or antagonisms (IPQ greater than 100%) are shown in Table 5. Forty seven % of all synergistic and antagonistic mixtures were rated with a higher risk-of-bias (i.e. mixtures judged with a “definitely high” and “probably high” risk-of-bias) compared to 38.5% of all additive mixtures. This suggests an association between the risk-of-bias scores and a non-additivity classification, although this was statistically non-significant (Fisher’s Exact test, alpha = 5%, Table 5).

3.5. Interactions by chemical class and study type (*in vitro*, *in vivo*, acute, chronic)

To investigate whether certain types of chemicals give rise to more frequent deviations from additivity, we stratified the entries in the *reappraisal database* according to type of chemicals and major use classes into mixtures containing metals, pesticides/biocides, endocrine disruptors, pharmaceuticals, and mycotoxins. We found that synergisms were slightly more frequent with mixtures composed of only pesticides and biocides. However, this trend was not statistically significant (chi-square test with Yates correction, Table 6). The corresponding IPQ distributions revealed no indications for chemical-class specific patterns, the medians were always close to zero (i.e. perfect additivity, Table 6).

Most mixture studies in the *reappraisal database* (>80%) employed only 2 or 3 components. This made it difficult to conclusively analyse an association between the occurrence of toxicological interactions and the number of mixture compounds. From the 78 mixtures that were classified as synergistic, ca. 78% (N = 61) were composed of 2 or 3 compounds.

As shown in Table 7, additivity appeared more likely to occur in mixtures tested in *in vivo* bioassays (N = 189, 66.3%) than *in vitro* test systems (N = 60, 60%), however the difference was not statistically significant and we therefore dismissed this observation as a chance finding. In *in vitro* studies synergisms were found more frequently than antagonisms. Among *in vivo* studies, the proportions of synergism and antagonisms were similar.

3.6. Specific concerns with synergisms

Warne and Hawker (1995) did not observe synergisms that exceeded 3-fold deviations from predicted additivity, and the synergisms identified by Boobis et al. (2011) were not more than 4-fold lower than predicted effect doses. Deviations of these magnitudes translate into IPQ values of −200% and −300%, respectively. We therefore looked for examples of synergistic interactions in our *reappraisal database* that exceed an IPQ of −300%. Table 8 presents the mixture experiments of

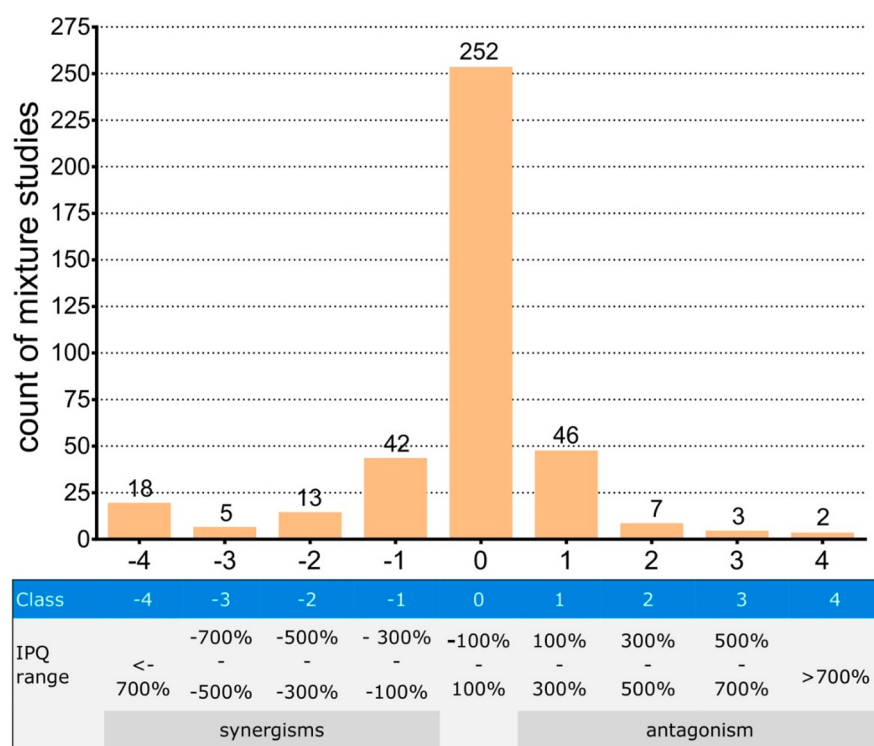


Fig. 7. Distribution of IPQ values (top) with definition of classes (bottom). Synergistic interactions are to the left, antagonisms to the right of the middle of the distribution.

Table 5

Interaction classification by risk-of-bias evaluation.

Interaction	Definitely High Risk	Probably High Risk	Probably Low Risk	Definitely Low Risk	Total
Synergistic	6 (20.0%)	30 (22.7%)	38 (17.5%)	4 (44.4%)	78 (20.0%)
Additive	19 (63.3%)	79 (59.9%)	150 (69.1%)	4 (44.4%)	252 (65.0%)
Antagonistic	5 (16.7%)	23 (17.4%)	29 (13.4%)	1 (11.2%)	58 (15.0%)
Total	30	132	217	9	388
Non-additive	High Risk 64 (47.1%)		Low Risk 72 (52.9%)		136 (100.0%)
Additive	98 (38.9%)		154 (61.1%)		252 (100.0%)
Total	162		226		388

human relevance that met our selection criterion, and Table 9 shows the experiments relevant to ecotoxicological endpoints.

Four of the ten experiments listed in Table 8 involve endocrine

Table 6

Interaction classification by chemical class.

Interaction	Only pesticides or biocides		Only endocrine disruptors		Only mycotoxins
	Only metals			Only pharmaceuticals	
Synergistic	13 (21.0%)	24 (27.9%)	5 (22.7%)	11 (22.2%)	3 (13.6%)
Additive	41 (66.1%)	49 (57.0%)	15 (68.2%)	30 (60.0%)	18 (81.8%)
Antagonistic	8 (12.9%)	13 (12.1%)	2 (9.1%)	9 (18.0%)	1 (4.6%)
Total	62 (100%)	86 (100%)	22 (100%)	50 (100%)	22 (100%)
IPQ					
Median	-11.1	-22.0	-18.1	15.5	-56.3
5%- 95% percentiles	-392 - 148.5	-900 - 202.5	-870-441	-771 - 153.4	-150-89.5

Table 7

Interaction classification by study type¹.

Interaction	In vitro	In vivo, acute	In vivo, chronic	Total
Synergistic	24 (24.0%)	41 (19.2%)	13 (18.1%)	78 (20.2%)
Additive	60 (60.0%)	145 (68.1%)	44 (61.1%)	249 (64.7%)
Antagonistic	16 (16.0%)	27 (12.7%)	15 (20.8%)	58 (15.1%)
Total	100	213	72	385

¹For three mixture studies a distinct classification into acute or chronic was not possible due to lack of reported details.

disruption endpoints relevant to androgen signalling and male sexual differentiation. The strongest synergism was observed in studies of suppressions of androgen receptor activation *in vitro* with a combination of 5 parabens (Kjærstad et al. 2010), with 100-fold lower mixture concentrations than anticipated based on concentration addition. In the same paper, mixtures of 3 azole fungicides were found to synergise in blocking androgen receptor activation, with 10-fold lower concentrations than predicted. Four-fold lower concentrations than expected were observed by Kjeldsen et al. (2013) with a combination of 5 pesticides. The molecular basis for these synergisms is unclear. Other studies of

Table 8
Studies of human toxicological relevance showing strong synergisms (IPQ < -300%).

Mixture ID ¹	Test Organism	Assay Name	Chemicals	CA [fold deviation]	IPQ [%]	Reference
14204-3	Chinese Hamster Ovary (CHO) cell line	<i>In vitro</i> AR reporter gene assay	methyl paraben; ethyl paraben;	0.01	-9900.00	(Kjærstad et al. 2010)
			propyl paraben; butyl paraben;			
175-1	Housefly (<i>Musca domestica</i>)	<i>In vitro</i> inhibition of AChE	iso-butyl paraben	0.02	-4900.00	(Arora and Kumar 2015)
174-1	Housefly (<i>Musca domestica</i>)	<i>In vitro</i> inhibition of AChE	monocrotophos;	0.04	-2400.00	(Arora et al. 2017)
14784-2	ACHN human renal adenocarcinoma cell line	<i>In vitro</i> cell viability assay	triazophos; carbofuran	0.05	-1900.00	(Choi et al. 2010)
			melamine; cyanuric acid			
2885-3	HepG2 liver cells	<i>In vitro</i> cytotoxicity assay	microcystin-LR; 17β-estradiol;	0.06	-1566.67	(Ma et al. 2017)
14204-4	Chinese Hamster Ovary (CHO) cell line	<i>In vitro</i> AR reporter gene assay	ractopamine	0.10	-900.00	(Kjærstad et al. 2010)
			epoxiconazole;			
			propiconazole;			
4437-2	Human lymphocytes	<i>In vitro</i> comet assay	tebuconazole	0.20	-403.90	(Sultana Shaik et al. 2016)
			chlorpyrifos; profenofos			
14784-3	Madin-Darby canine kidney (MDCK) epithelial cell line	<i>In vitro</i> cell viability assay	melamine; cyanuric acid	0.20	-400.00	(Choi et al. 2010)
2303-2	Chinese Hamster Ovary (CHO-K1) cell line	<i>In vitro</i> AR reporter gene assay	bitertanol;	0.25	-300.00	(Kjeldsen et al. 2013)
			propiconazole;			
			cypermethrin;			
			terbuthylazine;			
828-1	Wistar rats	<i>In vivo</i> reproduction/ developmental toxicity screening test	malathion	0.25	-300.00	(Christiansen et al. 2009)
			vinclozolin; finasteride;			
			DEHP; prochloraz			

¹ Mixture ID is as defined in the data extraction inventory (Supplementary Material S1) and the reappraisal database (Supplementary Material S4).

mixtures that included parabens in comparable experimental systems did not report deviations from additivity (Ermler et al. 2011; Orton et al. 2014).

The only *in vivo* study showing strong synergisms is by Christiansen et al. (2009). It involved a developmental toxicity model with the rat in which a combination of 4 chemicals capable of disrupting male sexual differentiation was shown to have synergistic interactions on malformations of the penis. Other androgen-sensitive endpoints analysed in the same animals with the same mixture (retained nipples, changes in anogenital distance) showed dose additive effects.

In a study examining the effects of binary mixtures of organophosphates on acetylcholinesterase (AChE) inhibition *in vitro*, Arora and Kumar (2015) and Arora et al. (2017) found strong synergisms (50-times lower concentrations than expected). Synergisms between organophosphates have previously been attributed to increases in the rate of activation to the AChE-inhibiting oxon forms by one or several compounds in the mixture. However, such interactions require P450-monoxygenase enzymes which were not present in the pure AChE preparations used by Arora and Kumar.

Other synergisms of note involve cytotoxicity produced by combinations of melamine and cyanuric acid (Choi et al. 2010) and microcystin-LR, 17β-estradiol and ractopamine (Ma et al. 2017). In addition, synergistic genotoxic responses were observed with a binary mixture of organophosphate pesticides (Sultana Shaik et al. 2016).

Among ecotoxicological studies (Table 9), strong synergisms were found with binary combinations of azole fungicides and pyrethroids on *Daphnia magna* (Bjergager et al. 2012; Nørgaard and Cedergreen 2010). These synergisms occurred at environmentally relevant concentrations (Bjergager et al. 2012).

Several studies documented strong synergisms between organophosphates and triazines and organophosphates and pyrethroids on zebrafish embryo mortality (Wang et al. 2017a). Synergisms between

organophosphates and pyrethroids are attributed to the ability of the organophosphate oxon form to inhibit esterases that inactivate pyrethroids. Triazines synergise with organophosphates by accelerating the formation of oxons through induction of P450 monooxygenases.

Some multi-component mixture experiments revealed strong synergisms (nearly 10-fold lower concentrations than expected), such as in the study by Petersen and Tollefsen (2012) of suppression of estrogen receptor-mediated vitellogenin induction in fish hepatocytes by combinations of 11 PAHs, PCBs and PCDDs. The higher than predicted anti-estrogenic potency of the mixture is ascribed to the ability of PAHs, PCBs and PCDDs to induce CYP1A1. CYP1A1 induction leads to a down-turn of estrogen receptor expression, thereby diminishing its activation, with larger than expected reductions in vitellogenin levels.

Chen et al. (2015) observed synergisms with combinations of 7 pesticides and metals on earthworm toxicity. Since the mixture contained several organophosphates, triazines and pyrethroids, the higher than expected toxicity can be traced to the established ability of these classes of compounds to synergise through metabolic interactions.

The ability of organophosphates to synergise with triazines and pyrethroids, and of azoles to interact with pyrethroids is well established and has been noted in earlier reviews (Cedergreen 2014). New evidence to strengthen these observations has emerged (e.g. Chen et al. 2015; Wang et al., 2017a). Several synergisms not observed previously with combinations of heavy metals have become apparent. Of note are Cr(VI) and cadmium, and nickel and cadmium with their stronger than expected effects on inhibiting algal growth (Mo et al. 2016).

We were unable to locate any new *in vivo* low dose studies that established synergistic interactions not already reviewed by Boobis et al. (2011).

Table 9
Studies of ecotoxicological relevance showing strong synergisms (IPQ < -300%).

Mixture ID ¹	Test Organism	Assay endpoint	Chemicals	CA [fold deviation]	IPQ [%]	Reference
175-1	Musca domestica (housefly)	<i>In vitro</i> inhibition of AChE -	monocrotophos; triazophos	0.02	-4900.00	(Arora and Kumar 2015)
5828-2	Daphnia magna (Water flea)	Immobilisation (under field conditions)	prochloraz; esfenvalerate	0.03	-3233.33	(Bjergager et al. 2012)
438-1	Pseudokirchneriella subcapitata (algae)	Growth inhibition	2,4,6-trichlorophenol;	0.03	-2841.18	(Boltes et al. 2012)
174-1	Musca domestica (housefly))	<i>In vitro</i> inhibition of AChE -	PFOS; triclosan carbofuran; triazophos	0.04	-2400.00	(Arora et al. 2017)
2904-1	Pseudokirchneriella subcapitata (algae)	Growth inhibition	cephalotine; ciprofloxacin	0.04	-2400.00	(Magdaleno et al. 2015)
4868-2	Bacillus subtilis (bacteria)	Microtox derivate	sulfamethoxypyridazine;	0.06	-1566.67	(Wang et al. 2018a)
8593-1	Daphnia Magna (Water flea)	Immobilisation	ormetoprim; oxytetracycline hydrochloride prochloraz; alpha-cypermethrin	0.08	-1150.00	(Nørgaard and Cedergreen 2010)
4971-1	Danio rerio (zebrafish)	Mortality (Fish Embryo)	atrazine; phoxim;	0.10	-900.00	(Wang et al. 2017a)
7274-1	Vibrio qinghaiensis spp. - Q67 (bacteria)	Microtox	lambda-cyhalothrin; aspirin; oxytetracycline	0.10	-900.00	(Mo et al. 2017)
12785-1	Vibrio phosphoreum (bacteria)	Microtox	malononitrile; acetaldehyde;	0.10	-900.00	(Li et al. 2014)
4967-1	Eisenia fetida (Earthworm)	Mortality (artificial soil assay)	p-nitrobenzaldehyde; benzaldehyde lambda-cyhalothrin; phoxim;	0.11	-809.99	(Wang et al. 2015)
703-1	Eisenia fetida (Earthworm)	Mortality	butachlor atrazine; phoxim; chlorpyrifos;	0.11	-809.99	(Chen et al. 2015)
3612-1	Rainbow trout hepatocytes	Vitellogenin production	cadmium; butachlor; imidacloprid; lambda-cyhalothrin; avermectin benz(a)anthracene;	0.12	-733.33	(Petersen and Tollefsen 2012)
14381-1	Vibrio phosphoreum (bacteria)	Microtox derivate	benzo(b)fluoranthene; benzo(a)pyrene; ZM 189.154; 4-hydroxytamoxifen; benzo(k)fluoranthene; indeno(1,2,3-cd)pyrene; PCB-77; PCB-126; 2,3,7,8-tetrachlorodibenzo-p-dioxin; beta-naphthoflavone	0.12	-733.33	(Tian et al. 2013)
3168-2	Selenastrum capricornutum (algae)	Growth inhibition	malononitrile; acetaldehyde Potassium dichromate;	0.13	-669.9	(Mo et al. 2016)
2801-1	Escherichia coli K-12 MG1655 (bacteria)	Growth inhibition	cadmium chloride 2.5-hydrate sulfadoxine; baquiloprim	0.14	-614.29	(Long et al. 2016)
5828-1	Daphnia magna (Water flea)	Immobilisation	prochloraz; esfenvalerate	0.15	-566.67	(Bjergager et al. 2012)
4868-1	Escherichia coli (bacteria)	Microtox	sulfamethoxypyridazine; ormetoprim	0.16	-525.00	(Wang et al. 2018a)
14830-1	Vibrio fischeri (bacteria)	Microtox	malononitrile; acetaldehyde;	0.16	-525.00	(Tian et al. 2012)
10497-2	Ankistrodesmus fusiformis (algae)	Growth inhibition	allyl alcohol; allyl chloride oxytetracycline; enrofloxacin	0.17	-488.24	(Carusso et al. 2018)
4971-2	Danio rerio (Zebrafish)	Mortality (Fish Embryo)	atrazine; phoxim; chlorpyrifos;	0.18	-455.56	(Wang et al. 2017a)
4970-1	Danio rerio (Zebrafish)	Mortality (Fish Embryo)	lambda-cyhalotrin kresoxim-methyl; cyprodinil	0.19	-426.32	(Wang et al. 2018b)
4973-1	Danio rerio (Zebrafish)	Mortality (Fish Embryo)	butachlor; imidacloprid;	0.19	-426.32	(Wang et al. 2017b)
1680-1	Cylindrotheca closterium (algae)	Growth inhibition	gamma-cyhalotrin tylosin; lincomycin; ciprofloxacin	0.20	-400.00	(Hagenbuch and Pinckney 2012)
3168-2	Chlorella pyrenoidosa (algae)	Growth inhibition	potassium dichromate;	0.20	-400.00	(Mo et al. 2016)
3168-3	Selenastrum capricornutum (algae)	Growth inhibition	cadmium chloride 2.5-hydrate nickel(II)nitrate hexahydrate;	0.20	-400.00	(Mo et al. 2016)
10037-1	RTG-2 cell line from Oncorhynchus mykiss (rainbow trout)	<i>In vitro</i> cytotoxicity	cadmium chloride 2.5-hydrate bisphenol A; diclofenac; naproxen; metoprolol; fluoxetine;	0.20	-400.00	(Fernández et al. 2013)

(continued on next page)

Table 9 (continued)

Mixture ID ¹	Test Organism	Assay endpoint	Chemicals	CA [fold deviation]	IPQ [%]	Reference
1280–2	Schizopera knabeni (Copepod)	Mortality	ketoprofen; caffeine; galaxolide; atenolol; propranolol; diphenylhydantoin; hydrochlorothiazide; cashmeran; ethinyl estradiol	0.22	–365.12	(Evans and Nipper 2007)
2829–10	Scenedesmus obliquus (algae)	Growth inhibition	2,4-dichloroaniline; beta-naphthol	0.24	–316.67	(Lu et al. 2007)
13449–1	Vibrio phosphoreum (bacteria)	Microtox	cadmium; p-benzenediol	0.24	–316.67	(Su et al. 2008)
1564–1	Pseudodiaptomus inopinus (planktonic copepod)	Mortality	tetrabromobisphenol A;	0.25	–300.00	(Gong et al. 2016)
14829–1	Zebrafish (Brachydanio rerio)	Mortality	hexabromocyclododecane phoxim; etofenprox	0.25	–300.00	(Zhang et al. 2010)

¹ Mixture ID is as defined in the data extraction inventory (Supplementary Material S1) and the reappraisal database (Supplementary Material S4).

4. Discussion

We identified many studies that describe deviations from predicted additivity, but in most cases these deviations, when quantitatively reappraised, were small. With the criteria proposed by Belden et al. (2007), most of these deviations were classed as additive.

With more than 50% of the 1,220 mixture experiments in our *data extraction inventory* assessed by their authors as synergistic, antagonistic or interactive, the proportion of claimed toxicological interactions is higher than in previous reviews where the share of mixtures exhibiting synergisms or antagonisms was between 10% (Deneer 2000), 12% (Belden et al. 2007), 7–26% (Cedergreen 2014) and 23% (Warne and Hawker 1995). However, quantitative comparisons between our effort and the earlier reviews should be made with caution because of differences in emphasis on certain groups of chemicals. The most likely explanation for our higher proportion of interactive studies lies in the importance we placed on data for quantitative reappraisal. In line with this review's objectives, we focused on outcomes suggesting toxicological interactions. When publications reported several mixture experiments, we selected those with the largest deviations from additivity. A more exhaustive data extraction strategy might have increased the proportion of reported additive outcomes.

Even though we deliberately focused on experiments which in their authors' opinions indicated strong interactions, our quantitative reappraisal of mixture experiments revealed that relatively few claims of synergisms or antagonisms fell outside the range of deviations which Belden et al. (2007) classed as additive. Nearly two thirds (65%) of author claims of interactions were re-evaluated as too small to be considered as synergistic or antagonistic. For the remainder of these studies (N = 136), we confirmed the authors' evaluations of interactions. This is equivalent to 11% of the 1220 experiments in our *data extraction inventory*, a proportion not too dissimilar to those reported in the reviews by Warne and Hawker (1995), Deneer (2000), Belden et al. (2007) and Cedergreen (2014).

Even so, these proportions must be judged with caution as they depend on the quantitative criterion chosen for the evaluation of deviations from predicted additivity. In the absence of statistical criteria, we applied the criteria proposed by Belden et al. (2007) and classed as interactions only IPQs outside the range between –100% and 100%. Application of the Broderius criteria for additivity (IPQs between –25% and 20%) would have classified a larger number of studies as interactive but would have paid insufficient regard to the experimental variations often encountered in *in vivo* studies. The coefficients of variation of 30% for intra- and inter-laboratory variations that OECD guidelines can consider as acceptable for demonstrating transferability of assays (OECD 2012) exceed the range of IPQs between –25% and 20%. Thus, IPQs outside this range could justifiably also be interpreted as indicating poor reproducibility rather than synergisms or antagonisms. For this reason, we did not apply the Broderius criteria and instead based our evaluations on the IPQ range proposed by Belden et al. (2007).

The possibility remains that there are some overlooked toxicological interactions among the experiments that reported additive mixture effects which we did not reappraise. We also could not evaluate experiments investigating potentiations, where the combination of one active and one inactive component leads to exacerbations of effects. In addition, there is the fraction of mixture studies that could not be re-evaluated at all, for missing data or incomplete reporting. Finally, it was difficult to evaluate observed mixture effects against additivity expectations derived from IA. As IA normally predicts larger additive mixture effect doses than DA, a mixture effect evaluated as additive relative to DA may be synergistic in terms of IA. This may have biased our re-evaluation. However, we do not expect this bias to be large. The magnitude of prediction differences between DA and IA is driven by the number of mixture components, and can become large only with mixtures composed of more than 3–4 chemicals (Kortenkamp et al. 2012). As noted earlier, most mixture experiments employed only 2 or 3 components, where the prediction differences between DA and IA are small.

Taken together, our findings support the use of DA (CA) as the default concept for anticipating the combined effects of chemical unless there is specific evidence that interactions might be relevant. Therefore, to achieve a sufficient degree of protection, this strategy must be complemented by an awareness of the synergistic potential of specific classes of chemicals. This includes combinations of triazine, azole and pyrethroid pesticides at environmentally relevant doses and should be extended to certain endocrine disrupting chemicals and metal compounds such as chromium (VI) and nickel in combination with cadmium.

With the increase in mixture studies during the last decade, there is now a good empirical basis for understanding how chemicals work together to produce combined toxicity. However, the field appears to be mired in studying binary mixtures. Although the theoretical and practical concepts necessary for conducting and interpreting multi-component mixture experiments are established and verified, very few studies go beyond binary or tertiary mixtures. With some studies, we encountered difficulties with extracting relevant experimental data, due to omission of important details (e.g. mixture ratios) and insufficient dose–response analyses. Future work could elaborate guidelines for the publication of mixture experiments, based on our risk-of-bias tool. There is also a dearth of studies designed to investigate additivity for combinations of chemicals at low doses, or at environmentally relevant mixture ratios. It appears that the field is currently over-descriptive, repetitive, and under-theorised. It should move on to address real-world challenges.

Declaration of Competing Interest

Olwenn V. Martin is a member of the management board of the European Chemical Agency. All other authors declare no competing interests.

Acknowledgements

This work was funded by the European Commission, Directorate-General Joint Research Centre, Service Contract CCR.F.933992.X0 which is gratefully acknowledged. We acknowledge the input of Alicia Paini, Andrea N Richarz, Sander van der Linden, and Andrew Worth in defining the protocol. Richard Evans is thanked for developing the architecture of the reappraisal database.

Appendix A. Supplementary data

Supplementary data to this article can be found online at <https://doi.org/10.1016/j.envint.2020.106206>.

References

- Arora, S., Balotra, S., Pandey, G., Kumar, A., 2017. Binary combinations of organophosphorus and synthetic pyrethroids are more potent acetylcholinesterase inhibitors than organophosphorus and carbamate mixtures: an in vitro assessment. *Toxicol. Lett.* 268, 8–16. <https://doi.org/10.1016/j.toxlet.2016.12.009>.
- Arora, S., Kumar, A., 2015. Binary combinations of organophosphorus pesticides exhibit differential toxicity under oxidised and un-oxidised conditions. *Ecotoxicol. Environ. Saf.* 115, 93–100. <https://doi.org/10.1016/j.ecoenv.2015.01.003>.
- Belden, J.B., Gilliom, R.J., Lydy, M.J., 2007. How well can we predict the toxicity of pesticide mixtures to aquatic life? *Integr. Environ. Assess. Manag.* 3, 364–372. [https://doi.org/10.1897/1551-3793\(2007\)3\[364:hwcwp\]2.0.co;2](https://doi.org/10.1897/1551-3793(2007)3[364:hwcwp]2.0.co;2).
- Bjergager, M.B.A., Hanson, M.L., Solomon, K.R., Cedergreen, N., 2012. Synergy between prochloraz and esfenvalerate in *Daphnia magna* from acute and subchronic exposures in the laboratory and microcosms. *Aquat. Toxicol.* 110–111, 17–24. <https://doi.org/10.1016/j.aquatox.2011.12.001>.
- Boltes, K., Rosal, R., García-Calvo, E., 2012. Toxicity of mixtures of perfluorooctane sulphonic acid with chlorinated chemicals and lipid regulators. *Chemosphere* 86, 24–29. <https://doi.org/10.1016/j.chemosphere.2011.08.041>.
- Boobis, A., Budinsky, R., Collie, S., Crofton, K., Embry, M., Felter, S., et al., 2011. Critical analysis of literature on low-dose synergy for use in screening chemical mixtures for risk assessment. *Crit. Rev. Toxicol.* 41, 369–383. <https://doi.org/10.3109/10408444.2010.543655>.
- Bopp, S., Berggren, E., Kienzler, A., van der Linden, S., Worth, A., 2015. Scientific methodologies for the assessment of combined effects of chemicals – a survey and literature review; doi: 10.2788/093511.
- Broderius, S.J., Kahl, M.D., Hoglund, M.D., 1995. Use of joint toxic response to define the primary mode of toxic action for diverse industrial organic chemicals. *Environ. Toxicol. Chem.* 14, 1591–1605. <https://doi.org/10.1002/etc.5620140920>.
- Caruso, S., Juárez, A.B., Moreton, J., Magdaleno, A., 2018. Effects of three veterinary antibiotics and their binary mixtures on two green alga species. *Chemosph* 194, 821–827. <https://doi.org/10.1016/j.chemosphere.2017.12.047>.
- Cedergreen, N., 2014. Quantifying synergy: a systematic review of mixture toxicity studies within environmental toxicology. *PLoS ONE* 9. <https://doi.org/10.1371/journal.pone.0096580>.
- Chen, C., Wang, Y., Qian, Y., Zhao, X., Wang, Q., 2015. The synergistic toxicity of the multiple chemical mixtures: Implications for risk assessment in the terrestrial environment. *Environ. Int.* 77, 95–105. <https://doi.org/10.1016/j.envint.2015.01.014>.
- Choi, L., Kwak, M.Y., Kwak, E.H., Kim, D.H., Han, E.Y., Roh, T., et al., 2010. Comparative nephrotoxicity induced by melamine, cyanuric acid, or a mixture of both chemicals in either Sprague-Dawley rats or renal cell lines. *J. Toxicol. Environ. Heal - Part A Curr. Issues* 73, 1407–1419. <https://doi.org/10.1080/15287394.2010.511540>.
- Christiansen, S., Scholze, M., Dalggaard, M., Vinggaard, A.M., Axelstad, M., Kortenkamp, A., et al., 2009. Synergistic disruption of external male sex organ development by a mixture of four antiandrogens. *Environ. Health Perspect.* 117, 1839–1846. <https://doi.org/10.1289/ehp.0900689>.
- Deneer, J.W., 2000. Toxicity of mixtures of pesticides in aquatic systems. *Pest Manag. Sci.* 56, 516–520. [https://doi.org/10.1002/\(SICI\)1526-4998\(200006\)56:6<516::AID-PS163>3.0.CO;2-O](https://doi.org/10.1002/(SICI)1526-4998(200006)56:6<516::AID-PS163>3.0.CO;2-O).
- Drakvik, Elina, Altenburger, Rolf, Aoki, Yasunobu, Backhaus, Thomas, et al., 2020. Statement on advancing the assessment of chemical mixtures and their risks for human health and the environment. *Environ. Int.*, 105267 <https://doi.org/10.1016/j.envint.2019.105267>.
- Ermiler, S., Scholze, M., Kortenkamp, A., 2011. The suitability of concentration addition for predicting the effects of multi-component mixtures of up to 17 anti-androgens with varied structural features in an in vitro AR antagonist assay. *Toxicol. Appl. Pharmacol.* 257, 189–197.
- European Commission (EC), 2012. Communication from the Commission to the Council - The combination effects of chemicals. Chemical mixtures. COM(2012) 252 Final 10.
- Evans, A.D., Nipper, M., 2007. Toxicity of phenanthrene and lindane mixtures to marine invertebrates. *Environ. Toxicol.* 22, 495–501. <https://doi.org/10.1002/tox.20279>.
- Evans, R.M., Martin, O.V., Faust, M., Kortenkamp, A., 2016. Should the scope of human mixture risk assessment span legislative/regulatory silos for chemicals? *Sci. Total Environ.* 543, 757–764. <https://doi.org/10.1016/j.scitotenv.2015.10.162>.
- Fernández, C., Carbonell, G., Babín, M., 2013. Effects of individual and a mixture of pharmaceuticals and personal-care products on cytotoxicity, EROD activity and ROS production in a rainbow trout gonadal cell line (RTG-2). *J. Appl. Toxicol.* 33, 1203–1212. <https://doi.org/10.1002/jat.2752>.
- Gong, W., Zhu, L., Hao, Y., 2016. Lethal and sublethal toxicity comparison of BFRs to three marine planktonic copepods: effects on survival, metabolism and ingestion. *Z. Yan, ed. PLoS ONE* 11 (e0147790). <https://doi.org/10.1371/journal.pone.0147790>.
- Grenier, B., Oswald, I.P., 2011. Mycotoxin co-contamination of food and feed: Meta-analysis of publications describing toxicological interactions. *World Mycotoxin. J.* 4, 285–313. <https://doi.org/10.3920/WMJ2011.1281>.
- Hagenbuch, I.M., Pinckney, J.L., 2012. Toxic effect of the combined antibiotics ciprofloxacin, lincomycin, and tylosin on two species of marine diatoms. *Water Res.* 46, 5028–5036. <https://doi.org/10.1016/j.watres.2012.06.040>.
- Kienzler, A., Berggren, E., Bessems, J., Bopp, S., Van Der, L.S., Worth, A., 2014. Assessment of Mixtures. Rev. Regulatory Requirements Guidance 136. <https://doi.org/10.2788/138523>.
- Kjærstad, M.B., Taxvig, C., Andersen, H.R., Nellemann, C., 2010. Mixture effects of endocrine disrupting compounds in vitro. *Int. J. Androl.* 33, 425–433. <https://doi.org/10.1111/j.1365-2605.2009.01034.x>.
- Kjeldsen, L.S., Ghisari, M., Bonefeld-Jørgensen, E.C., 2013. Currently used pesticides and their mixtures affect the function of sex hormone receptors and aromatase enzyme activity. *Toxicol. Appl. Pharmacol.* 272, 453–464. <https://doi.org/10.1016/j.taap.2013.06.028>.
- Kortenkamp, A., Evans, R., Faust, M., Kalberlah, F., Scholze, M., Schuhmacher-Wolz, U., 2012. Investigation of the state of the science on combined actions of chemicals in food through dissimilar modes of action and proposal for science-based approach for performing related cumulative risk assessment; doi:10.2903/sp.efsa.2012.EN-232.
- Kortenkamp, Andreas, Faust, Michael, 2018. Regulate to reduce chemical mixture risk - regulatory systems must better provide for risks from exposure to multiple chemicals. *Science* 361 (6399), 224–225. <https://doi.org/10.1126/science.aat9219>.
- Kousta, E., Lam, J., Sutton, P., Johnson, P.I., Atchley, D., Sen, S., et al., 2013. Applying the navigation guide: case study #1. the impact of developmental exposure to perfluorooctanoic acid (PFOA) on fetal growth. A Systematic Review of the Non-Human Evidence: Final Protocol. Available: https://prhe.ucsf.edu/sites/g/files/tksra341/f/PFOA_NON-HUMAN_PROTOCOL.pdf [accessed 15 May 2020].
- Li, W., Tian, D., Lin, Z., Wang, D., Yu, H., 2014. Study on the variation rules of the joint effects for multicomponent mixtures containing cyanogenic toxicants and aldehydes based on the transition state theory. *J. Hazard. Mater.* 267, 98–108. <https://doi.org/10.1016/j.jhazmat.2013.12.035>.
- Long, X., Wang, D., Lin, Z., Qin, M., Song, C., Liu, Y., 2016. The mixture toxicity of environmental contaminants containing sulfonamides and other antibiotics in *Escherichia coli*: Differences in both the special target proteins of individual chemicals and their effective combined concentration. *Chemosphere* 158, 193–203. <https://doi.org/10.1016/j.chemosphere.2016.05.048>.
- Lu, G., Wang, C., Tang, Z., Guo, X., 2007. Joint toxicity of aromatic compounds to algae and QSAR study. *Ecotoxicology* 16, 485–490. <https://doi.org/10.1007/s10646-007-0153-1>.
- Ma, M., Pi, F., Wang, J., Ji, J., Sun, X., 2017. New insights into cytotoxicity induced by microcystin-LR, estradiol, and ractopamine with mathematical models: individual and combined effects. *Chemosphere* 168, 223–233. <https://doi.org/10.1016/j.chemosphere.2016.10.004>.
- Magdaleno, A., Saenz, M.E., Juárez, A.B., Moreton, J., 2015. Effects of six antibiotics and their binary mixtures on growth of *Pseudokirchneriella subcapitata*. *Ecotoxicol. Environ. Saf.* 113, 72–78. <https://doi.org/10.1016/j.ecoenv.2014.11.021>.
- Marking, L.L., Mauck, W.L., 1975. Toxicity of paired mixtures of candidate forest insecticides to rainbow trout. *Bull. Environ. Contam. Toxicol.* 13, 518–523. <https://doi.org/10.1007/BF01685172>.
- Martin, O.V., Bopp, S., Ermiler, S., Kienzler, A., McPhie, J., Paini, A., et al., 2018. Protocol for a systematic review of ten years of research on interactions in chemical mixtures of environmental pollutants. *Zenodo*, 10.5281/ZENODO.1319759.
- Mo, L., Zheng, M., Qin, M., Zhang, X., Liu, J., Qin, L., et al., 2016. Quantitative characterization of the toxicities of Cd-Ni and Cd-Cr binary mixtures using combination index method. *Z. Ren, ed Biomed Res. Int.* 2016 (4158451) <https://doi.org/10.1155/2016/4158451>.
- Mo, L.Y., Liu, J., Qin, L.T., Zeng, H.H., Liang, Y.P., 2017. Two-stage prediction on effects of mixtures containing phenolic compounds and heavy metals on vibrio qinghaiensis sp. Q67. *Bull. Environ. Contam. Toxicol.* 99 (17–22) <https://doi.org/10.1007/s00128-017-2099-1>.
- Nørgaard, K.B., Cedergreen, N., 2010. Pesticide cocktails can interact synergistically on aquatic crustaceans. *Environ. Sci. Pollut. Res.* 17, 957–967. <https://doi.org/10.1007/s11356-009-0284-4>.
- OECD, 2012. Series on Testing and Assessment: No. 179, Validation Report (Phase 2) for the Zebrafish Embryo Toxicity Test. ENV/JM/MONO(2012)25.
- Orton, F., Ermiler, S., Kugathas, S., Rosivatz, E., Scholze, M., Kortenkamp, A., 2014. Mixture effects at very low doses with combinations of anti-androgenic pesticides, antioxidants, industrial pollutant and chemicals used in personal care products. *Toxicol. Appl. Pharmacol.* 278, 201–208.
- Petersen, K., Tollefsen, K.E., 2012. Combined effects of oestrogen receptor antagonists on in vitro vitellogenesis. *Aquat. Toxicol.* 112–113, 46–53. <https://doi.org/10.1016/j.aquatox.2012.01.023>.
- Shamseer, L., Moher, D., Clarke, M., Ghersi, D., Liberati, A., Petticrew, M., et al., 2015. Preferred reporting items for systematic review and meta-analysis protocols (prisma-p) 2015: elaboration and explanation. *BMJ* 349. <https://doi.org/10.1136/bmj.g7647>.
- Su, L., Yuan, X., Mu, C., Yan, J., Zhao, Y., 2008. Evaluation and QSAR study of joint toxicity of substituted phenols and cadmium to photobacterium phosphoreum. *Chem. Res. Chinese Univ.* 24, 281–284. [https://doi.org/10.1016/S1005-9040\(08\)60060-4](https://doi.org/10.1016/S1005-9040(08)60060-4).

- Sultana Shaik, A., Shaik, A.P., Jamil, K., Alsaeed, A.H., 2016. Evaluation of cytotoxicity and genotoxicity of pesticide mixtures on lymphocytes. *Toxicol. Mech. Methods* 26, 588–594. <https://doi.org/10.1080/15376516.2016.1218577>.
- Tian, D., Lin, Z., Ding, J., Yin, D., Zhang, Y., 2012. Application of the similarity parameter (λ) to prediction of the joint effects of nonequotoxic mixtures. *Arch. Environ. Contam. Toxicol.* 62, 195–209. <https://doi.org/10.1007/s00244-011-9695-6>.
- Tian, D., Lin, Z., Yin, D., 2013. Quantitative structure activity relationships (QSAR) for binary mixtures at non-equitoxic ratios based on toxic ratios-effects curves. *Dose-Response* 11, 255–269. <https://doi.org/10.2203/dose-response.11-042.Lin>.
- Vijver, M.G., Elliott, E.G., Peijnenburg, W.J.G.M., de Snoo, G.R., 2011. Response predictions for organisms water-exposed to metal mixtures: A meta-analysis. *Environ. Toxicol. Chem.* 30, 1482–1487. <https://doi.org/10.1002/etc.499>.
- Wang, D., Wu, X., Lin, Z., Ding, Y., 2018a. A comparative study on the binary and ternary mixture toxicity of antibiotics towards three bacteria based on QSAR investigation. *Environ. Res.* 162, 127–134. <https://doi.org/10.1016/j.envres.2017.12.015>.
- Wang, Y., Chen, C., Qian, Y., Zhao, X., Wang, Q., 2015. Ternary toxicological interactions of insecticides, herbicides, and a heavy metal on the earthworm *Eisenia fetida*. *J. Hazard. Mater.* 284, 233–240. <https://doi.org/10.1016/j.jhazmat.2014.11.017>.
- Wang, Y., Dai, D., Yu, Y., Yang, G., Shen, W., Wang, Q., et al., 2018b. Evaluation of joint effects of cyprodinil and kresoxim-methyl on zebrafish, *Danio rerio*. *J. Hazard. Mater.* 352, 80–91. <https://doi.org/10.1016/j.jhazmat.2018.03.023>.
- Wang, Y., Lv, L., Yu, Y., Yang, G., Xu, Z., Wang, Q., et al., 2017a. Single and joint toxic effects of five selected pesticides on the early life stages of zebrafish (*Danio rerio*). *Chemosphere* 170, 61–67. <https://doi.org/10.1016/j.chemosphere.2016.12.025>.
- Wang, Y., Yang, G., Dai, D., Xu, Z., Cai, L., Wang, Q., et al., 2017b. Individual and mixture effects of five agricultural pesticides on zebrafish (*Danio rerio*) larvae. *Environ. Sci. Pollut. Res.* 24, 4528–4536. <https://doi.org/10.1007/s11356-016-8205-9>.
- Warne, M.S., Hawker, D.W., 1995. The number of components in a mixture determines whether synergistic and antagonistic or additive toxicity predominate: the funnel hypothesis. *Ecotoxicol. Environ. Saf.* 31, 23–28. <https://doi.org/10.1006/eesa.1995.1039>.
- Whaley, P., Aiassa, E., Beausoleil, C., Beronius, A., Bilotta, G., Boobis, A., et al., 2020. A Code of Practice for the Conduct of Systematic Reviews in Toxicology and Environmental Health Research (COSTER). *Zenodo*.
- Zhang, Z.Y., Yu, X.Y., Wang, D.L., Yan, H.J., Liu, X.J., 2010. Acute toxicity to zebrafish of two organophosphates and four pyrethroids and their binary mixtures. *Pest Manag. Sci.* 66, 84–89. <https://doi.org/10.1002/ps.1834>.