Examining the feasibility of mixture risk assessment: A case study using a tiered approach with data of 67 pesticides from the Joint FAO/WHO Meeting on Pesticide Residues (JMPR)

Richard M. Evans[*], Martin Scholze, Andreas Kortenkamp

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

ARTICLE INFO

ABSTRACT

The way in which mixture risk assessment (MRA) should be included in chemical risk assessment is a current topic of debate. We used data from 67 recent pesticide evaluations to build a case study using Hazard Index calculations to form risk estimates in a tiered MRA approach in line with a Framework proposed by WHO/IPCS. The case study is used to illustrate the approach and to add detail to the existing Framework, and includes many more chemicals than previous case studies.

A low-tier MRA identified risk as being greater than acceptable, but refining risk estimates in higher tiers was not possible due to data requirements not being readily met. Our analysis identifies data requirements, which typically expand dramatically in higher tiers, as being the likely cause for an MRA to fail in many realistic cases. This forms a major obstacle to routine implementation of MRA and shows the need for systematic generation and collection of toxicological data. In low tiers, hazard quotient inspection identifies chemicals that contribute most to the HI value and thus require attention if further refinement is needed. Implementing MRA requires consensus on issues such as scope setting, criteria for performing refinement, and decision criteria for actions.

© 2015 Elsevier Ltd. All rights reserved.

1. Introduction

Mixture risk assessment (MRA) is the assessment of the cumulative risk to human health or the environment from multiple chemicals via multiple routes. Currently, chemicals are routinely assessed on a chemical-by-chemical basis, with the notable exception of the approach to dioxin-like chemicals, wherein selected PCBs, dioxins and furans are assessed collectively by application of a toxic equivalency quotient/factor (TEQ/TEF) approach (van den Berg et al., 1998). There is concern that the chemical-by-chemical approach may not be sufficiently protective if two or more chemicals have the same toxic effect (Boobis et al., 2008; Kortenkamp et al., 2009). It is incontrovertible that humans are exposed to more than one chemical at a time, for example to the multiple chemicals found in food, in air and drinking water, and in household and consumer products and cosmetics. Mixture toxicology is the branch of toxicology that deals with predicting and managing the exposure of humans or the environment to multiple chemicals and their associated toxicological effects. The existence of a mixture *per se* does not always indicate a risk to human or environmental health, but indicates the need to examine whether more accurate estimations of risk will be produced by considering all of the chemicals that are present.

Whilst there is a broad consensus on the basic science of mixture toxicology (Kortenkamp et al., 2009; DG Health and Consumer Protection, 2011), the path to regulatory implementation of these considerations, as an MRA, in chemical risk assessment is less clear. Options were outlined in an EFSA opinion (EFSA,

2008) and, currently, proposals for MRA approaches include a Framework developed by WHO/IPCS for "Risk assessment of combined exposure to multiple chemicals" (Meek et al., 2011), a decision tree of the European Commission Scientific Committees (DG Health and Consumer Protection, 2011) and an approach examining the contribution of individual mixture components to the joint effect, termed maximum cumulative ratio (Price et al., 2014). Of these, the WHO/IPCS Framework is the most widely used. It has the stated aim of aiding "risk assessors in identifying priorities for risk management for a wide range of applications where co-exposures to multiple chemicals are expected." The Framework is described as hierarchical, phased and tiered with "integrated and iterative consideration of exposure and hazard at all phases, with each tier being more refined". A 'more refined' tier is described as being less cautious, more certain, more labour intensive and more data intensive than the preceding tier. The underlying philosophy is to invest more resources in the analysis only if assessments based on less data intensive assumptions indicate that levels deemed to be acceptable are exceeded. The tiers detailed in the WHO/IPCS Framework are not fixed; their use will depend on data availability, and tiers can be added or removed as necessary. Use of either predictive or probabilistic methodologies is placed in various tiers and uncertainty is considered at each tier.

Two areas in which the WHO/IPCS Framework does not provide much detail are 1) criteria for ceasing refinement and applying risk management measures, and 2) criteria for the grouping of chemicals within an MRA. A decision about ceasing refinement is needed at the end of each tier when the risk has not been shown to be acceptable. It is not clear whether the implementation of risk management that would be mandated if the highest tier was exceeded should also be mandated in low tiers when progression is not achievable due to data gaps or difficulties with data availability. Grouping of chemicals for MRA is proposed in the second tier of the WHO/IPCS Framework but no details are provided on what the need or prerequisites for grouping are. Outstanding questions include, for example, would grouping on the basis of chemical structure be acceptable? Should the grouping approach have particular demands in terms of retaining conservatism or would the Framework allow this property to be lost? EFSA has begun the process of identifying cumulative assessment groups (CAGs), commencing with the definition of CAGs covering phenomenological effects of pesticides on thyroid and nervous system (EFSA, 2013), although the full set of CAGs may need to be available before they can meaningfully be introduced into MRA. We have explored options in both these areas within this case study.

The guiding approach that is used in most MRA approaches is the Hazard Index (HI), in which firstly, hazard quotients (HQs) are calculated for each chemical in the exposure scenario by dividing their exposure level by an 'acceptable' level, such as an acceptable daily intake (ADI) or reference dose (RfD); secondly, the HQs are summed to give the HI (Teuschler and Hertzberg, 1995). Conventionally, a HI of greater than one indicates that the total exposure exceeds the level considered to be 'acceptable', where the definition of acceptable depends on the denominators used in the HQ calculation. The Margin of Exposure (MoE) approach is conceptually similar to the HI, but usually operates with 'points of departure' (PoD) values such as benchmark doses or no-observed adverse effect levels (NOAELs) to which safety or uncertainty factors have not been applied. Whereas the critical value for an HI is *greater than or equal to* 1, the critical value for an MoE is *less than or equal to* one hundred.

Two prior case studies have presented examples of MRA for triazole pesticides and for polybrominated diphenyl ethers (PBDEs) (EFSA, 2009; Meek et al., 2011). The triazole case study used the hazard index (HI) approach to explore a tiered strategy in detail, but artificially restricted their analysis to seven or eleven pesticides for endpoints of cranio-facial malformation and hepatotoxicity, respectively, for reasons of data availability (EFSA, 2009). The study calculated low tier HI values that were mostly below one: 0.1 (total Dutch population) and 0.24 (children sub-population). However, when HI values were calculated for individual food commodities, as part of an evaluation of the use of HI in maximum residue level (MRL) setting, exposure to bitertanol via apples had a HI of 1.19, which reduced to 0.17 in the next tier.

The PBDE case study dealt with a complex situation comprising seven components, each of which was itself a mixture of PBDE congeners. A Tier 1 assessment produced MoEs of 300 (based on upper-bound of a deterministic exposure estimate) or approximately 30 (based on biomonitoring) and, despite 30 being below the critical value of 100 for MoEs, the authors considered that the in-depth evaluation of human health risks from PBDE mixtures was a 'low priority' (Meek et al., 2011). Further refinement in higher tiers and the need for risk management was not explored. Both studies showed the need for further case studies to explore the possible outcomes for scenarios that are different to those reported so far; and in this paper we provide a scenario involving many more chemicals (67) than have been previously considered.

Here, we present a case study based on a dataset compiled from evaluations of 67 pesticides by the Joint FAO/WHO Meeting on Pesticide Residues (JMPR) between 2006 and 2010. We use this case study to explore the options for refinement within the hazard portion of a tiered MRA approach following the conceptual approach of the WHO/IPCS Framework (Meek et al., 2011). Our aim was to use a relatively large, regulatory data set to explore how refinement options affect the outcome of MRA where differing amounts of data are available. We have utilized the international estimated daily intake (IEDI) values calculated for 67 pesticides in annual JMPR reports from 2006 to 2010 and use this dataset to work through the tiers of the proposed Framework. The case study does not represent an actual MRA for pesticides, rather it is used to understand and explore the tiered approach and to explore the consequences of differing data requirements and assumptions for MRA.

2. Materials and methods

The guiding approach used in this case study is the Hazard Index (HI), which is calculated using the formula:

$$HI = \sum_{i=1}^{n} \frac{EL_i}{AL_i}$$

where *EL* is the exposure level, *AL* is the acceptable level, and *n* is the number of chemicals in the mixture. A hazard quotient (HQ) is calculated for each chemical, by dividing *EL* by *AL*, and the HQs are summed to give the HI. Various measures for exposure levels and acceptable levels may be applied; the only constraint is that both must be expressed in the same unit. Input values for AL can be, for example, Acceptable Daily Intake (ADI) values or Reference Doses (RfD) for specific endpoints. Where mean values are given, the ordinary arithmetic mean was used.

The dataset used in this case study was compiled from exposure and risk data provided for 67 pesticides that were evaluated in the five annual Joint FAO/WHO Meeting on Pesticide Residues (JMPR) reports from 2006 to 2010 (JMPR, 2010; JMPR, 2009; JMPR, 2008; JMPR, 2007; JMPR, 2006). 76 evaluations were included, with 9 pesticides being evaluated twice. JMPR reports establish acceptable daily intakes (ADIs) and also report international estimated daily intakes (IEDIs) which are calculated on a weight per person basis

for 13 cluster diets in the Global Environment Monitoring System-Food contamination and assessment programme (GEMS/Food, http://www.who.int/foodsafety/areas_work/chemical-risks/gems-food/en/). Cluster diets represent sets of countries grouped together on the basis of food consumption patterns. The 178 countries making up each cluster diet are listed in Table SI1. IEDIs are derived by multiplying the concentrations of residues (supervised trials median residues (STMRs), supervised trials median residue in processed commodity (STMR-Ps) or recommended MRLs) by the average daily per *caput* consumption estimated for each commodity on the basis of the GEMS/Food diet. Prior to 2006, the GEMS cluster diets were organised into 5 regions rather than 13, and hence data prior to 2006 is not readily compatible with data from 2006 onwards and was not included.

When IEDI values are expressed as percentages of the ADI (as in Annex 3 of each JMPR report), they are equivalent to the hazard quotients (HQ) required for the calculation of a hazard index (HI). The compiled data is presented in Table SI2 and has been scaled so that the HQ and HI are on a unitary scale instead of percentage; on the unitary scale, values greater than 1 indicate a greater than acceptable risk for the single chemical (for HQ values) or mixture (for HI values). HQ values are given to 3 decimal places, so that values less than 0.001 are considered effectively zero. ADIs, and the toxicological data used in their derivation, were collated from individual toxicological evaluations provided by JMPR for each of the pesticides (Table SI3). Three of the evaluations referred to more than one chemical by including an isomer or related structure. We have referred to these as if they are single chemicals throughout the case study, as follows (text used by JMPR given in brackets): Cyfluthrin ('CYFLUTHRIN with BETA-CYFLUTHRIN'), Cyhalothrin ('CYHALOTHRIN (including Lambdacyhalothrin)'), Triadimefon ('TRIADIMEFON with TRIADIMENOL'). The JMPR used two region-specific bodyweights of 60 Kg (regions A to F, H to K and M) and 55 Kg (regions G and L); and these region-specific body weights were used for calculations throughout this case study.

Chemical structures for each pesticide were retrieved from the PPDB (Pesticide Properties DataBase, http://sitem.herts.ac.uk/aeru/footprint/index2.htm) as Simplified Molecular Input Line Entry Specification (SMILES) codes and ToxTree software (version 2.5) was used to assign chemicals to the appropriate Cramer class based on their SMILES code. All of the chemicals were members of Cramer class III, which receives a Threshold of Toxicological Concern (TTC) value of 90 μg/person per day (all regions except G and L where TTC of 82.5 μg/person is used due to an assumed average bodyweight of 55 kg instead of 60 kg). The PPDB was also used as the source for health effect profiles for each pesticide.

3. Results

The full dataset used in this case study is provided in Table SI2, which lists the 67 pesticides and 13 geographical regions included in the study, and presents HQ and HI calculations. It is notable that the HQ of an individual chemical exceeded 1 in only two instances, both for chlorpyrifos-methyl, in regions C and H in 2009 (JMPR, 2009). In nine instances the same pesticide was evaluated twice within the 2006–2010 period, and in all cases the second evaluation resulted in a higher IEDI than the previous evaluation (Figure SI1). The increase ranged from 11% (Difenoconazole, 2007/2010) to 404% (Buprofezin, 2008/2009).

3.1. Tier 1: hazard index (HI) using ADI values

The results of a hazard index (HI) analysis are presented in Fig. 1 and Table SI2. The table shows HQs for 67 chemicals for 13 GEMS diet regions. The HI (sum of all HQs) was greater than one for all
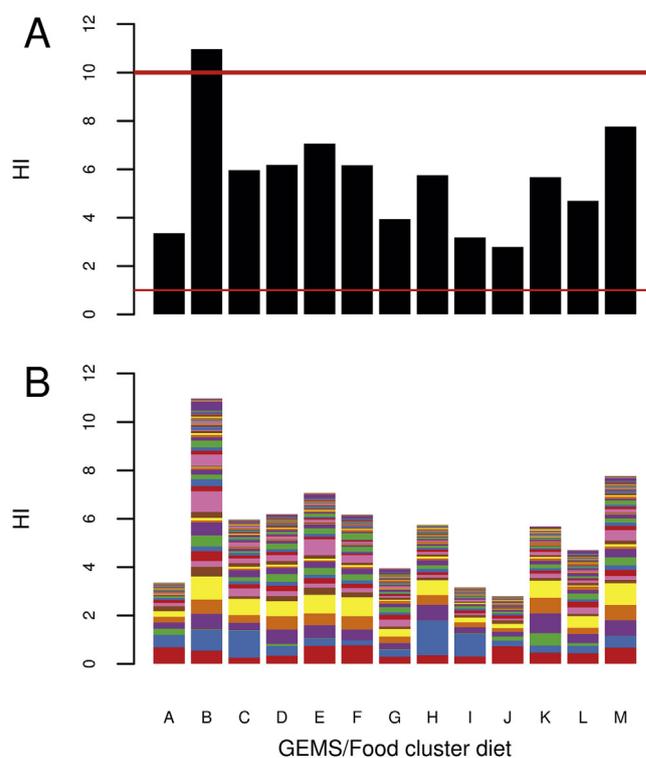

Fig. 1. HI values for thirteen GEMS diet regions in Tier 1 MRA. A) bar chart showing HI values for thirteen GEMS diet regions, thin horizontal red line indicates the critical value of 1, thick horizontal red line indicates a HI of 10. **B)** bar chart showing each HI value segmented into HQ values; HQ values are sorted within each bar using the decreasing order of magnitude for region A so that coloured segments in different bars represent the same chemical. HQ values for particular chemicals can be distinguished in Fig. 2 and Table SI2. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

regions, and exceeded 10 in one region (Fig. 1, Table SI2). HIs ranged from 2.8 (Region J, Africa) to 11 (Region B, Africa/Europe/Middle East). In no case was the HI solely driven by a single chemical exceeding a HQ of 1; even in two instances when a single HQ exceeded 1, removal of these values would not have reduced the HI to less than one.

Individual values for each GEMS diet region are presented in Fig. 1A and in Table SI2 (final row). In all regions, the HI exceeded the critical value of one, indicating that an acceptable level of risk cannot be concluded and that refinement of the assessment is required. In one case, the HI was larger than 10 (region B), indicating a calculated risk that is more than ten times greater than the critical value (of one). At this tier, the calculation of HI values greater than one means that a conclusion that risk is acceptable cannot be reached, and indicates the need for further refinement by moving to a more data intensive analysis.

The way in which HQ values for each chemical contribute to the overall HI value is shown graphically in Figs. 1B and 2. HQ values ranged from a maximum of 1.4 down to effectively zero (Table SI2). Fig. 2 shows that the HQ distribution is clearly skewed for all regions, indicating that not all chemicals contribute equally to the HI. In Fig. 1B the HI value for each region is broken down into striped segments representing the individual HQ values that cumulate to the HI value and this reveals that HQ values for particular chemicals varied between the different regions. For example the red segment closest to the x-axis in each bar (Fig. 1B) indicates the HQ of fenitrothion which had the highest HQ in region A but made a proportionately smaller contributions in regions such as C, G and I. Fig. 2 shows cumulative distributions of HQ values for each region,
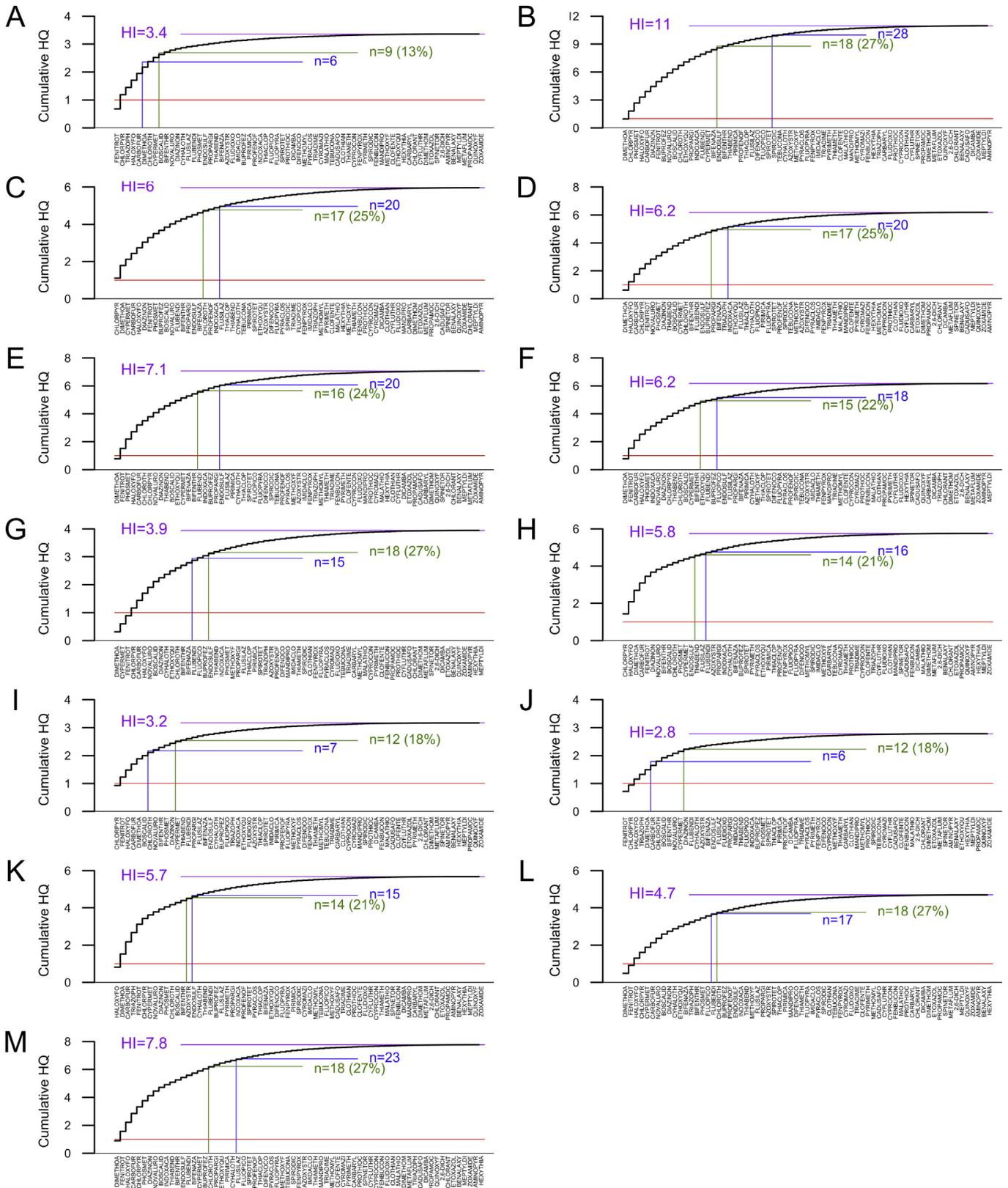
Fig. 2. HQ cumulative distributions in Tier 1. Graphs show a cumulative distribution of HQ values sorted in decreasing order of magnitude for each GEMS/Food region, indicated by the panel letters A–M. Pesticide names are given along the x-axis but are more easily read by zooming in on the online figures or from the tables provided (Table SI2). For ease of viewing, each graph is replicated at one graph per page as Figure SI2. Due to space constraints, only the first 8 characters of each name is given in these figure. Horizontal lines indicate the critical value of one (red line) and the HI (purple line). The interpretation of these values is described in more detail in the text. The HI value is shown on each graph (purple text); as is the number of chemicals (n) that, when ranked by decreasing contribution to the HI, cause the HI to exceed one (blue lines and text); and the number (n) and percentage of the chemicals that drive 80% of the HI (green lines and text). (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)
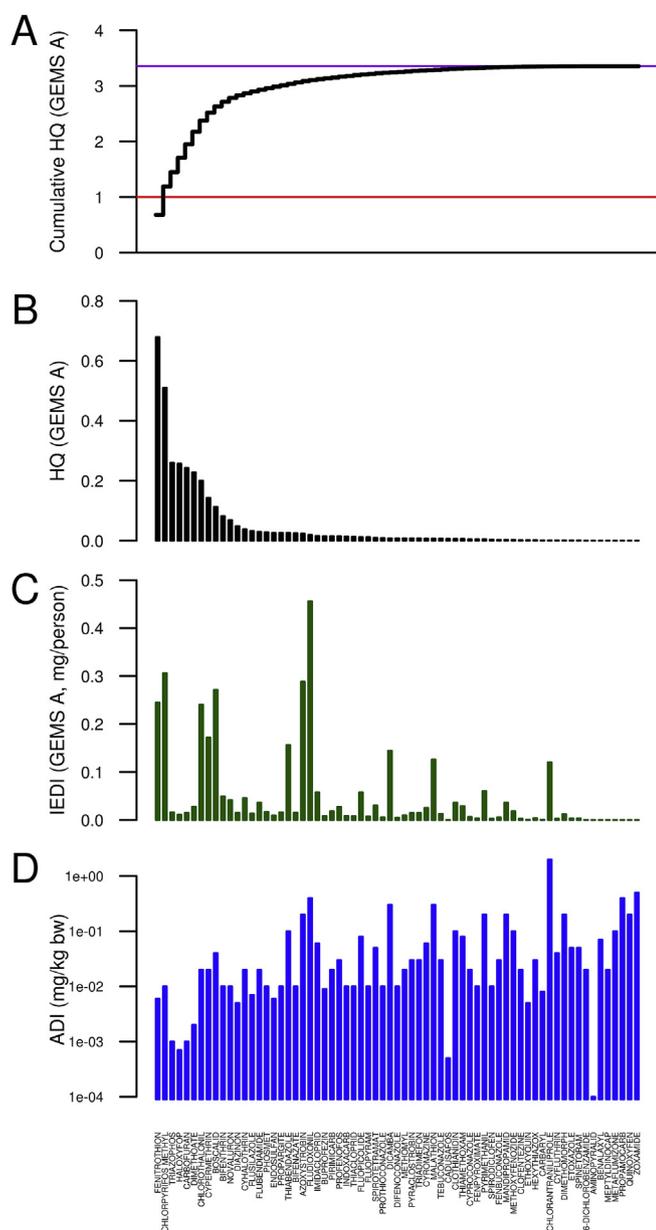
Fig. 3. Comparison of ADI, exposure and Tier 1 HQ distributions for region A. Bar charts showing: **A)** Cumulative distribution of HQ values (IEDI/ADI expressed in common units); **B)** HQ values; **C)** international estimated daily intakes (IEDI, mg/person); **D)** ADI (JMPR, mg/kg bw). Data in A-C is for GEMS region A. Bars are sorted in decreasing size of HQ (panel B).

Fig. 4. Comparison of HI with other summary metrics. Bar charts show different metrics that express particular aspects of the mixture assessment for each GEMS cluster diet and compares each metric to the HI. Metrics shown are A: HI, shown for reference; B: the largest HQ in each diet ("Max. HQ"); C: the number of chemicals that would cause the HI to equal or exceed one, when the actual HQs observed are ranked in descending order ("N_{HI-1}(ranked HQs)"); D: the mean HQ (ordinary arithmetic mean); E: the number of chemicals that would cause the HI to equal or exceed one, assuming each chemical has the mean HQ ("N_{HI-1}(mean HQ)"). Horizontal red lines indicates the critical value of 1 (thin line, A, B), 10 (thick line, A) and 1 divided by the total number of chemicals, i.e. 1/67 (D). Scatter plots on the right of bar charts B–E show the correlation of the metric shown in the bar chart (y-axis) with the HI (x-axis). (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

and non-cumulative distributions are also provided in Figure SI4. Fig. 2 shows that in two regions (C, H) the HI exceeded one when there was only a single chemical contributing, whilst region G showed the most gradual 'ramp' of the distribution, in that 4 chemicals are included before the cumulative HQ sum exceeds one. Fig. 2 also provides an indicator of how many chemicals are responsible for the bulk (set at 80%) of the HI, which ranged from 9 (13% of the 67 chemicals included) to 18 (27%). To allow easier comparison of the HQ distributions, all thirteen distributions were sorted (in decreasing order of HQ magnitude), normalized to the HI for each region and superimposed in one figure (Figure SI3). This graph reinforces the observation that the bulk of these HI values (80%, dotted horizontal line in Figure SI3) is derived from a sizeable minority of chemicals (9–18 chemicals, 13–27%), when the contributions are sorted in decreasing magnitude.
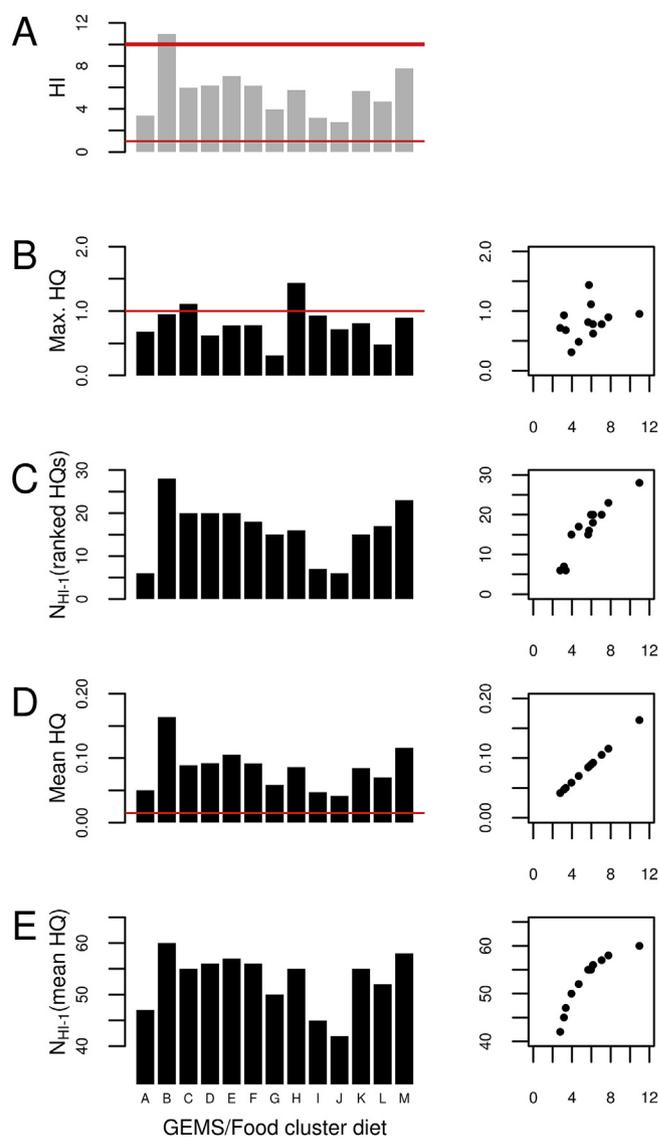
Each HQ value derives from two factors: an acceptable level (which was the ADI in this case study) and an exposure level (IEDI). Therefore we examined whether either factor drives the HI more than the other. Fig. 3 shows the distribution of HQ values for region A (as an example) and the underlying exposure level (IEDI) and acceptable level (ADI) for the set of pesticides in this case study. All distributions are plotted in order of descending HQ value, and the figure shows that neither exposure levels nor acceptable levels mirrored the resulting HQ distribution, indicating that the HQ distribution is generally driven by a combination of the underlying ADI and exposure distributions, and not directly by either one.

3.1.1. Summary metrics

The most information rich way of presenting the HI result in an MRA is as a graph of the cumulative distribution of HQ values (e.g. Fig. 2), however simple ways to describe or summarise the mixture risk, for example as a single number, may be useful when there is a need to compare many assessments, when viewing and comparing many detailed graphs is impractical. Fig. 4 shows several other metrics that can be readily derived from HIs: the largest HQ per cluster diet (Fig. 4B); the number of chemicals whose HQs, when ranked in descending order, add up to the portion of the HI that is above 1 ("N_{HI-1}(ranked HQs)", Fig. 4C), i.e. if these HQs could be refined or managed such that they become zero then the remaining chemicals would have a HI below 1; the average HQ (ordinary arithmetic mean, Fig. 4D); and the number (N) of chemicals that will cause the HI to exceed 1 if the average HQ is used ("N_{HI-1}(mean HQ)", Fig. 4E). No single metric is fully descriptive by itself, depending on the need to know about the size of the overall risk estimate, the largest single driver of the effect or the extent to which the risk estimate is driven by multiple chemicals (i.e. whether the overall risk does indeed come from a 'mixture'). Fig. 4 shows that most of the metrics do not correlate exactly with the HI, indicating that they indeed provide another dimension to that provided in the HI (which chiefly indicates the magnitude of the overall predicted risk without providing any indication of the underlying distributions of values).

Fig. 4B shows that the largest HQ in any diet was 1.4, for Chlorpyrifos-methyl in region H (Latin America). Chlorpyrifos-methyl also contributed the highest HQ in region C (Africa/Middle East; 1.1) and region I (Africa; 0.93). The next highest HQs approached, but did not exceed, 1 and were due to dimethoate in region B (Africa/Europe/Middle East; 0.952) and region M (Europe/Latin America; 0.897). This metric indicates which chemical is most responsible for the size of the HI; however it does not indicate that all the mixture 'risk' that is present comes from that one chemical. In this study even if the largest HQ in each diet could be completely removed, none of the HI values would fall below 1.

Fig. 4C shows N_{HI-1}(ranked HQs), the number of chemicals that constitute the portion of the HI that is above 1; this metric is also shown on each graph in Fig. 2 (blue lines). N_{HI-1}(ranked HQs) ranged from 6 (diets A and J) to 28 (diet B), and shows that, in the best case, if the chemicals leading to the six largest HQs could be eliminated then the remaining 61 chemicals would have an HI of just below 1, which could be deemed acceptable. In the worst case, 28 chemicals would need to be eliminated.

Fig. 4D shows a graph of average HQs, which exactly mirrors the HI graph (Fig. 4A) because the same number of chemicals was included for each cluster diet. However the critical value, indicated by a red line at 1/67 (that is, the critical value for a HI, divided by the number of chemicals included), shows that - in all of the cluster diets - the size of the HQ for the average chemical will cause the HI to exceed one, which is indeed what is seen for the HI (Figs. 1A and 4A). Fig. 4E shows that when the mean HQ is used to calculate the metric N_{HI-1}(mean HQ), values ranged from 42 for diet J (Africa) up to 60 for diet B (Africa). This metric can be compared to N_{HI-1}(ranked HQs), shown in Fig. 4C, which is a similar calculation but uses the observed HQs ranked in descending order rather than the mean HQ. For example, for diet A, N_{HI-1}(ranked HQs) was 6 chemicals but N_{HI-1}(mean HQ) was 48 chemicals, 8-fold more. The difference was not as large in other regions, for example in region B N_{HI-1}(ranked HQs) was 28 chemicals and N_{HI-1}(mean HQs) was 60 chemicals, 2.1-fold more. N_{HI-1}(mean HQ) was always larger than N_{HI-1}(ranked HQs), because the HQ distributions are clearly skewed (e.g. see Fig. 2), and in such situations averages are not accurate representations of the actual distributions. This observation indicates that a low tier analysis will be insightful for any mixture

scenario in which the HQ distribution is skewed, and could be useful for assessing the extent of any mixture problem. It then gives an indication of the effects likely to be required i.e. the resources available for refinement or risk management can be focused on the subset of chemicals which make a disproportionately large contribution to the HI.

3.2. Tier 0: HI analysis using TTC values (pseudo-tier 0)

In this case study, a Tier 1 analysis based on ADIs was possible because all of the included chemicals have been assigned ADIs. However in some scenarios it is likely that MRA will need to include one or more chemicals that have not yet been assigned an ADI, and the TTC has been proposed for use in such cases. We used our case study, in which all ADI were known, to explore likely outcomes for scenarios when one or more ADI are lacking. If a chemical has not been assigned an ADI it has been suggested that the TTC approach could be used in a tier 'zero' MRA, for example as explored by Price et al. (2009). We have therefore applied the TTC approach to the data set used in this case study to explore its impact. We note that the TTC concept is not intended for use when chemicals have been assigned ADIs, as was the case here, however we are using the approach here to assess its impact rather than to perform a risk assessment. We have also tested the extreme case when TTC values are used for all of the chemicals in the assessment.

An HI assessment using TTC values, HI (TTC), is presented in Figure SI5 which compares the HQ and HI values calculated using ADIs (Tier 1) or TTC values (pseudo-Tier 0). HI (TTC) values ranged from 37.5 to 146 and were up to 16 times greater than HI (ADI)s for the same regions (Figure SI5). For example the HI (ADI) for region A was 3.36 whilst the HI (TTC) for the same region was 37.6 (Figure SI5). The TTC might typically be applied to only a few components in a mixture assessment, and the impact of the use of TTC values should be assessed whenever they are used because risk estimates driven by the use of TTC values are clear candidates for refinement before risk management action is taken. In this case study, the mean HQ (TTC) value within each GEMS diet region ranged from 0.6 to 2.2, and the mean over all GEMS regions was 1.1, indicating that if a TTC value was applied to a single chemical, while the remainder are assessed using their ADIs, then on average the HI would be expected to exceed one in all cases. HQ (TTC) values actually ranged from 0.001 to 16 and the impact of smaller values would not greatly affect an overall HI, whilst larger values could be fully responsible for a risk estimate exceeding 1. The distribution of these values would only apply generally if other chemicals follow similar risk and exposure distributions to the pesticides used in this case study.

3.3. Higher tier analyses

Given that HI values calculated in Tier 1 exceed one, we attempted to conduct a higher tier assessment. We compiled the information used in the individual risk assessment of each pesticide by JMPR. Table SI3 lists the NOAEL value for each chemical and supporting information, including safety factor values. The most common safety factor (SF) was 100 (applied to 59 out of 67 chemicals, 88%). The data in Table SI3 shows that the NOAEL for each chemical was derived from a different endpoint, as identified by JMPR. In consequence, higher tier MRA requires knowledge of the highest dose tested that is without observable effect for all 67 endpoints and for each chemical. The reason for this high data requirement is that, although the ADI for each chemical was set on the basis of a different primary effect, all 67 chemicals could, in theory, cause a common additional effect, at doses only slightly above the NOAEL used in ADI setting, and this effect could cumulate

and would have a risk only slightly lower than that of the HI calculated from ADIs. In fact, this tier would require information for all of the included chemicals on all of the endpoints that might be of concern, not limited to endpoints on which any ADI was based. This is a substantial data requirement, both in the experimental work required and the effort involved in compilation of the data. It was not feasible to attempt this data collection for this case study, since the data, if available, are contained in dossiers in free text or in tables, released only in reports with an often inconsistent structure, and published in pdf format. None of these features facilitate data extraction and compilation.

Since collection of the data required to perform a higher tier was not feasible, this MRA would have to cease at this stage, with low tiers having failed to conclude an acceptable risk, because HI values exceeded one, and with this tier being unable to complete due to data requirements. The data requirement of this stage is sufficiently high that many assessments would not be completed at this stage, however this does not affect MRAs in which it is possible to conclude acceptable risk in low tiers, as there is no need to proceed to higher tiers in such cases. In the next section, we illustrate a possible refinement step using a surrogate dataset in place of the full set of toxicological information.

3.4. Tier 2: refinement of HI values using effect data

In order to illustrate how an MRA might progress in higher tiers, we used a surrogate dataset of human health classifications from the Pesticide Properties database (PPDB) for each of the 67 pesticides. The PPDB includes nine human health issues, which are listed in Table SI4 and shown in Fig. 5. The PPDB health issues are not regulatory categories and may not be suitable for regulatory purpose but they are used here to illustrate the approach. In addition,

some of the effects are acute, for which comparison to the IEDI (a chronic measure) could result in an underestimation of risk. For the purpose of this case study it is assumed that the PPDB health issues cover all of the effects of interest, and that they are assigned reliably. The PPDB assigns one of four classifications for each chemical for each health issue as follows: known to cause an effect; known not to cause the effect; data unclear; or data unavailable. The classifications from PPDB for all 67 pesticides included in this case study are shown in Fig. 5 and refined HI values are shown in Fig. 6. Note that, at this tier, there is now a HI value for each included endpoint (Fig. 6).

Using the PPDB classification it is possible to exclude the HQ of a chemical from the HI calculation if the chemical is known not to cause the effect under consideration (green segments, Figure SI7). It would not be sufficiently conservative to only include the chemicals that are known to cause an effect, since in many cases there are numerous chemicals for which the data are classified as unclear or unavailable. Consequently all three situations (known, unclear and data unavailable) are included in Fig. 6 but their relative contributions are shown as coloured segments of the bar representing each HI so that it can be seen how much of the revised HI is due to risks that are known or to risks arising from conservative assumptions (i.e. assuming a contribution to the joint effect unless proven otherwise). If quantitative data is available for each endpoint, for example a PoD, then a refined HI can be calculated using reference values for each endpoint, rather than repeatedly using the value for the most sensitive endpoint (as is implied by using the ADI). However, calculating refined HIs was not possible in this example.

In general, this approach would reduce risk estimates when chemicals can reasonably be excluded from consideration, for example, when they are known not to cause a given effect. In this case study, if only the 'known' risks are considered (red portions, Fig. 6) some of the health effects show a HI below one. However when the risks for chemicals that have unclear or unavailable data (blue and grey portions, Fig. 6) are also included, none of the HI assessments are below one. Thus, in this case study, it was not possible to draw a conclusion of acceptable risk at any of the tiers considered.

4. Discussion

A tiered MRA approach was used for 67 pesticides across thirteen GEMS diet regions, and produced low tier HI values that exceeded one in all regions, meaning that acceptable risk could not be concluded at this tier. Further refinement, for example by grouping according to common health effects was not possible due to high data requirements, however the approach that could be used was illustrated using PPDB health issue classifications. In one region, the HI value was greater than 10, demonstrating the need for guidance on interpreting the magnitude of the HI. When the HI exceeds one in low tiers, visualizing the HQ distribution may be helpful in showing which chemicals have a significant impact on the MRA — these are optimal candidates for refinement of either the exposure or hazard assessment in a higher tier. If many chemicals have a significant impact, then this provides an indication that the resources required to refine, and presumably reduce, risk estimates in higher tiers may be substantial. Conversely, resources do not need to be expended on refining either risk or hazard data for chemicals that do not contribute significantly.

In this study, the HQ of an individual chemical exceeded one on two occasions; both for chlorpyrifos-methyl in 2009, in regions C and H (JMPR, 2009). The JMPR has stated that "Percentages above 100 [equivalent to a HQ above 1] should not necessarily be interpreted as giving rise to a health concern because of the conservative
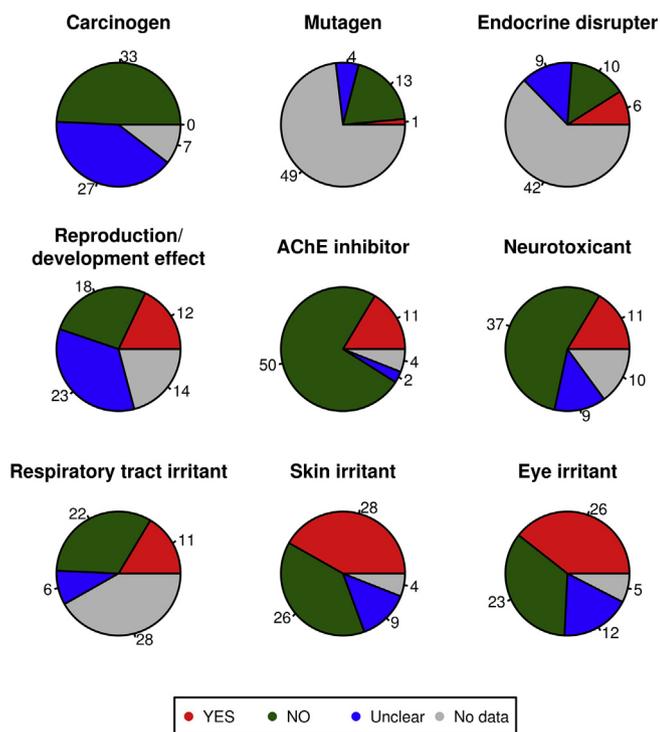

Fig. 5. Data confidence distributions for nine health effects as assigned in PPDB. Pie-charts showing the data confidence assigned by PPDB to the pesticides included in this case study for each of nine health issues (PPDB). Data confidence was assigned for each pesticide as 'yes': known to cause a problem; 'no': known not to cause a problem; '?': possibly, status not identified; or 'nd': no data.
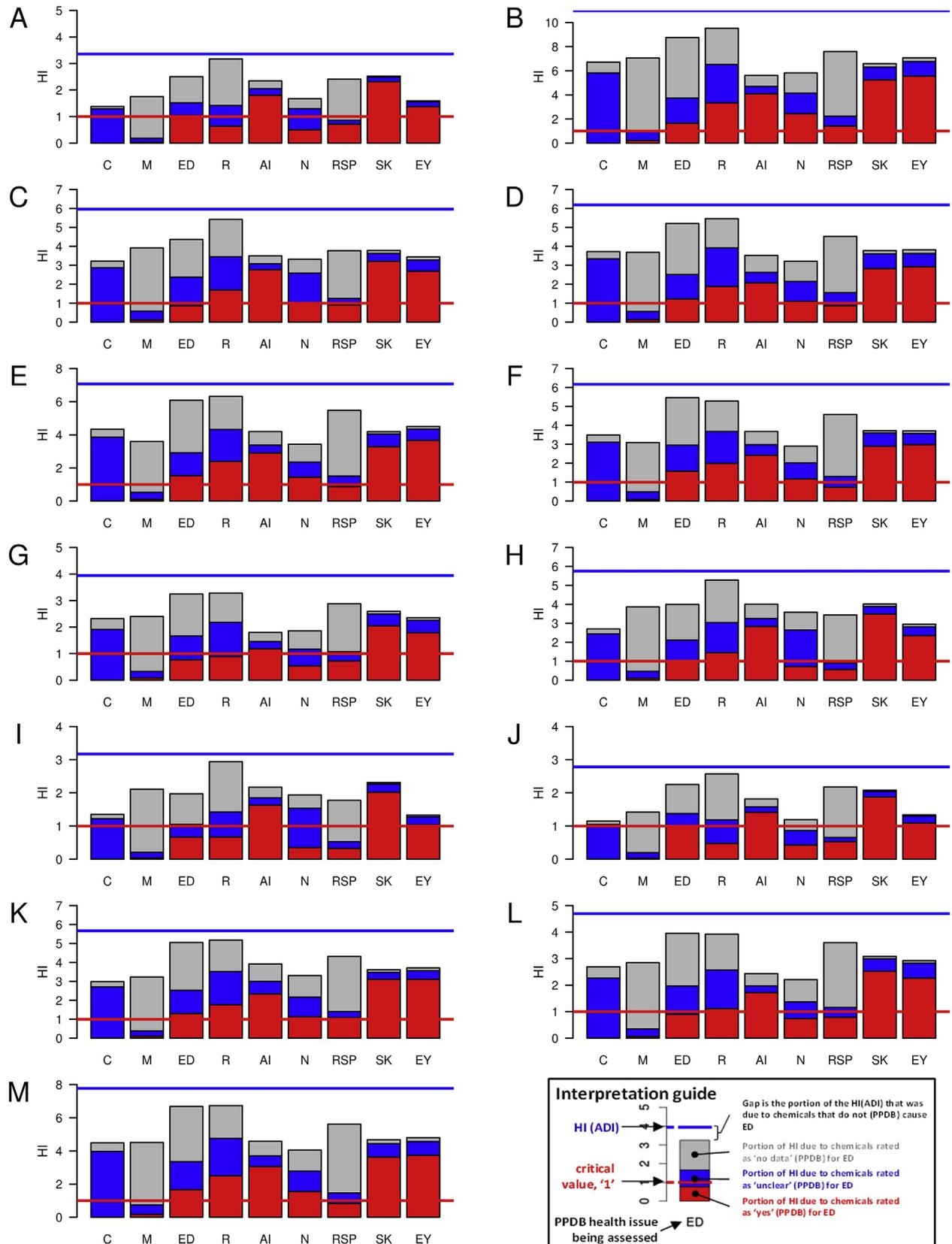
Fig. 6. HI values adjusted in Tier 2 MRA using PPDB classifications. Bar charts showing the HI for each health effect and each GEMS region (indicated by letters A-M) segmented by the data confidence assigned to each health issue and for each chemical. The graphic at bottom right provides a visual guide to interpretation of the graphs. Colours as in Fig. 5. Health issues are abbreviated as C, carcinogen; M, mutagen; ED, endocrine disrupter; R, reproduction/developmental effect; AI, acetylcholinesterase inhibitor; N, neurotoxicant; RSP, respiratory tract irritant; SK, skin irritant; EY, eye irritant. Horizontal blue lines indicate the HI (ADI) value calculated in Tier 1. Horizontal red lines indicate the critical value of one. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

assumptions used in the assessments." (JMPR, 2009). However, we note that the assumptions being used have been criticized and may be inappropriate for MRA, not least because they were not developed for this purpose (Martin et al., 2013). When a single HQ exceeds one then the HI will necessarily exceed one, but the scenario should still be examined because other components may nonetheless contribute significantly to the mixture risk. In any case there is clearly a need for robust guidance on the appropriate level of 'trigger' values for HQ and HI and on decision criteria for action accompanying such triggers, for example the resort to risk management when further refinement has not been possible.

The HI is a composite of hazard and exposure data for the included chemicals, however the overall driver of the HI value would appear to be the number of chemicals included. Opinions vary as to what this number would be in a realistic human scenario, and this constitutes an important knowledge gap. For example, in this case study if the average HQ is assumed, the HI would equal or exceed one if the number of 'average' chemicals was more than 6 (region B) or 24 (region J), in the best and worst case, respectively. It might be unlikely that all 67 of the chemicals in this case study, or even 24 of them, would share a common effect but it does not seem implausible that 6 or 7 chemicals might have a common effect, and this would be enough to generate a HI > 1 given the exposure pattern of for example, in this case study, region B.

In a tiered assessment, there are 3 outcomes at each tier: 1) risk is acceptable = MRA stops; 2) risk is not acceptable, refinement is possible = proceed to next tier; 3) risk is not acceptable, refinement is not possible = MRA stalls. Risk management should be considered either when higher tiers are not available or when MRA stalls, and there is a need to consider whether the high data demands of higher tiers make this likely to happen frequently, particularly when many chemicals are present in the MRA. The availability of a precautionary, low tier is valuable in allowing low risk scenarios to be rapidly identified without the expenditure of significant resource.

The issue of data availability and accessibility also applies to the use of mode or mechanism of action information. As well as there being no agreed framework to structure these discussions, the data that would be required to consider mode/mechanism is typically not available, is not a regulatory requirement, and would require significant resources to generate (Carmichael et al., 2011). As a result we have proposed that such considerations should not be included in the low tiers of MRA frameworks, as there would be a strong likelihood of the data requirement not being met, causing the MRA to stall; instead we proposed grouping only on the basis of common toxicological effect (Kortenkamp et al., 2012). In this way, completion of the MRA at lower tiers with relatively low data demands has the ability to avoid the need for detailed analyses in higher tiers if the conservative, somewhat worse-case, approaches used in low tiers already indicate low risk. The proposed approach rewards data rich situations by allowing assessments to be refined in higher tiers when accurate, derived values can be used in place of default, sometime worse-case, assumptions. Higher (more refined) tiers should be defined so that they cannot produce a higher estimate of risk (as this would indicate that the lower tiers were not precautionary enough) and will usually, but not always, result in a lower estimate of risk.

Although the purpose of this case study was not to perform an actual MRA for human pesticide exposure, we briefly consider whether these results are relevant to human health risks from pesticides. Due to the low tier nature of the calculations, and because our theoretical implementation of an MRA strategy failed before higher tiers were reached, we can only conclude that our analysis does not indicate any risk but also provides no evidence for the absence of risk. Our assessment includes many uncertainties, such as: IEDI values were cumulated without knowledge of the correlations between exposures to different pesticides; simultaneous exposure to all 67 pesticides was assumed; GEMS diets are set for a wide geographical area and there may be significant regional variation within each diet, which we could not consider; low tier HI calculation assumes a worst case that the ADI values are for a common endpoint shared by all the components and this could be shown to not be the case in this study (Table SI3); only 67 pesticides were included, not all pesticides in current use, and non-pesticides, such as dietary components, pollutants, additives and pharmaceuticals were not included; only dietary exposure was considered, not other routes.

This case study identifies several generic issues affecting the implementation of MRA: to reduce the impediments to performing MRA, toxicological data need to be accessible and collated across endpoints and across chemicals. The data for hazard and exposure should be in comparable metrics and available in an open, standardised format, and data summarization and censoring should be avoided. The need for consistency for the purposes of mixture assessment may be in tension with the flexibility that is possible in the single chemical risk assessment process, and this tension should be managed so that both single chemical and mixture assessments are well served. In order to be useful for mixture assessments, chemical testing should not be predicated on the most sensitive endpoint because, although this is appropriate for single chemical assessment, it becomes an obstacle to the assessment of multiple chemicals which do not have the most sensitive effect in common. It follows, that an MRA will only be able to consider those endpoints for which data are available for most of the included chemicals. The generation of new data for a particular endpoint may be so impractical as to be virtually impossible if the separate testing of many chemicals would be required. Finally, an MRA should start with a scoping step that identifies the chemicals to be included and the likely data availability so that achievable tiers can be constructed, as there is little value in attempting tiers that cannot be completed due to data requirements. Trigger values for HI and decision criteria for each tier should be set at the scoping stage.

In conclusion, our case study shows that single chemical data are not currently sufficient for use in MRA; that the number of chemicals likely to be present in a mixture scenario is a major knowledge gap; and that there is a need for clear, science-based guidance on decisions resulting from HI calculation at different tiers, including the explicit need for risk management action if either the highest tier is exceeded or if higher tiers are needed but cannot be completed due to data gaps. Other major issues in the regulatory implementation of MRA include the need for 1) consideration of the human body as a single receiving point for chemicals; 2) inclusion of all relevant chemicals and routes in MRA without artificial restriction; and 3) better understanding of whether the assumptions used in single chemical assessments are suitable or protective when used in a mixture context.

Acknowledgements

Appendix A. Supplementary data

Supplementary data related to this article can be found at http://dx.doi.org/10.1016/j.fct.2015.08.015.

Transparency document

Transparency document related to this article can be found online at http://dx.doi.org/10.1016/j.fct.2015.08.015.

References

Boobis, A.R., Ossendorp, B.C., Banasiak, U., Hamey, P.Y., Sebestyen, I., Moretto, A., 2008. Cumulative risk assessment of pesticide residues in food. Toxicol. Lett. 180, 137–150.

Carmichael, N., Bausen, M., Boobis, A.R., Cohen, S.M., Embry, M., Fruijtier-Polloth, C., Greim, H., Lewis, R., Bette Meek, M.E., Mellor, H., Vickers, C., Doe, J., 2011. Using mode of action information to improve regulatory decision-making: an ECE-TOC/ILSI RF/HESI workshop overview. Crit. Rev. Toxicol. 41, 175–186.

DG Health and Consumer Protection, 2011. Toxicity and Assessment of Chemical Mixtures. Opinion of the Scientific Committee on Consumer Safety, Scientific Committee on Health and Environmental Risks, Scientific Committee on Emerging and Newly Identified Health Risks.

EFSA, 2008. Opinion of the Scientific Panel on Plant Protection Products and Their Residues to Evaluate the Suitability of Existing Methodologies and, if Appropriate, the Identification of New Approaches to Assess Cumulative and Synergistic Risks from Pesticides to Human Health with a View to Set MRLs for Those Pesticides in the Frame of Regulation (EC) 396/2005.

EFSA, 2009. Scientific Opinion on Risk Assessment for a Selected Group of Pesticides from the Triazole Group to Test Possible Methodologies to Assess Cumulative Effects from Exposure through Food from These Pesticides on Human Health.

EFSA, 2013. Scientific Opinion on the Identification of Pesticides to be Included in Cumulative Assessment Groups on the Basis of Their Toxicological Profile.

JMPR, 2006. Pesticide Residues in Food 2006; Joint FAO/WHO Meeting on Pesticide Residues.

JMPR, 2007. Pesticide Residues in Food 2007; Joint FAO/WHO Meeting on Pesticide Residues.

JMPR, 2008. Pesticide Residues in Food 2008; Joint FAO/WHO Meeting on Pesticide Residues. Report 2008.

JMPR, 2009. Pesticide Residues in Food 2009; Joint FAO/WHO Meeting on Pesticide Residues. Report 2009.

JMPR, 2010. Pesticide Residues in Food 2010; Joint FAO/WHO Meeting on Pesticide Residues. Report 2010.

Kortenkamp, A., Backhaus, T., Faust, M., 2009. State of the Art Report on Mixture Toxicity.

Kortenkamp, A., Evans, R.M., 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 (An External Scientific Report for EFSA).

Martin, O.V., Martin, S., Kortenkamp, A., 2013. Dispelling urban myths about default uncertainty factors in chemical risk assessment—sufficient protection against mixture effects? Environ. Health 12, 53.

Meek, M.E., Boobis, A.R., Crofton, K.M., Heinemeyer, G., Raaij, M.V., Vickers, C., 2011. Risk assessment of combined exposure to multiple chemicals: a WHO/IPCS framework. Regul. Toxicol. Pharmacol. 60 (Suppl. 2). S1–S14.

Price, P., Zaleski, R., Hollnagel, H., Ketelslegers, H., Han, X., 2014. Assessing the safety of co-exposure to food packaging migrants in food and water using the maximum cumulative ratio and an established decision tree. Food Addit. Contam. Part A Chem. Anal. Control Expo. Risk Assess. 31, 414–421.

Price, P.S., Hollnagel, H.M., Zabik, J.M., 2009. Characterizing the noncancer toxicity of mixtures using concepts from the TTC and quantitative models of uncertainty in mixture toxicity. Risk Anal. 29, 1534–1548.

Teuschler, L.K., Hertzberg, R.C., 1995. Current and future risk assessment guidelines, policy, and methods development for chemical mixtures. Toxicology 105, 137–144.

van den Berg, M., Birnbaum, L., Bosveld, A.T., Brunstrom, B., Cook, P., Feeley, M., Giesy, J.P., Hanberg, A., Hasegawa, R., Kennedy, S.W., Kubiak, T., Larsen, J.C., van Leeuwen, F.X., Liem, A.K., Nolt, C., Peterson, R.E., Poellinger, L., Safe, S., Schrenk, D., Tillitt, D., Tysklind, M., Younes, M., Waern, F., Zacharewski, T., 1998. Toxic equivalency factors (TEFs) for PCBs, PCDDs, PCDFs for humans and wildlife. Environ. Health Perspect. 106, 775–792.