The development of new toxicity testing and approval processes for oil spill treatment products in the UK

Helen E Walton [1], Joshua J Davison [1], Joanna Uzyczak [1], Christopher Martin [1], Paula Milliken [1], Mark F Kirby [1]

[1]Centre for Environment, Fisheries and Aquaculture Science, Pakefield Road, Lowestoft, NR33 0HT, United Kingdom

ABSTRACT

Abstract Number: 687653

Current experimental protocols for the toxicity assessment of oil spill treatment products in the UK have been established since the 1970s. To address health and safety, cost and scientific robustness issues the UK approach for dispersant testing and approval has been reviewed and updated for implementation during 2020. To provide more robust scientific advice for the risk assessments that enable effective decision making on the use of oil remediation products in the event of a spill there has been a focus on methods that already have internationally accepted protocols. Standardisation of dispersant testing will promote more effective cross-institute comparisons of toxicity data and will enable further harmonisation of approaches in the future.

It is preferable that environmentally relevant test species are used but, as the scientific literature provides little conclusive evidence of a taxa-specific trend in sensitivity, species selection based on sensitivity alone was not justified. Eight dispersants, commonly stockpiled in the UK, were tested independently and in combination with a representative crude oil (Kuwait). Testing of dispersants in combination with oil has historically provided more variable results so this study has considered the benefits of this versus product only testing.

Core test species included the harpacticoid copepod, *Tisbe battagliai*, and the algae, *Skeletonema* sp., as both have cost-effective internationally standardised methods, whilst also being environmentally representative and using test species easily cultured under laboratory conditions with no seasonality. Other candidate test species, such as oyster embryos, had limitations in applicability due to seasonal issues. Fish testing was not considered as there was no ethical reasoning for vertebrate testing due to the absence of taxa-specific toxicity.

Results showed that, if oil is excluded from the assessment, *Skeletonema* sp. and *Tisbe battagliai*, can produce reliable, reproduceable and interpretable results. When running the *T. battagliai* test, independently on multiple occasions, without oil, dispersant 1, 2 and 3 had EC_{50} results that were not statistically different. This suggests that product only testing is suitable for ranking products based on toxicological hazard.

The redevelopment of the UK guideline to use standardised testing and the selection of appropriate, environmentally relevant test organisms will increase the quality and reliability of data used to underpin the UK oil spill treatment testing and approval scheme. The adoption of this approach will enable an approved list of products for use in UK waters to be maintained. However, the decision for dispersant use in any given scenario will need to be underpinned through expert advice applying a risk assessment approach taking account a range of incident-specific physical and environmental sensitivity information.

INTRODUCTION

Although the number of significant oil spills from tankers is declining on a global basis, the seaborne oil trade has grown steadily from 1970 to the present day (ITOPF, 2019) and the increase in tanker (and other vessel) movements poses a greater risk that major marine pollution incidents occur. The primary purposes of using oil spill treatment products are to prevent an oil slick from reaching the coastline, impacting sensitive biota that reside on the surface or

shore e.g. seabirds, marine mammals, to promote natural degradation and to reduce risks to spill responders by reducing Volatile Organic Compound (VOC) exposures to air. However, while the major toxicological issue is the oil, spill treatment products (e.g. dispersants) have been shown to have an inherent toxicity of their own and have the potential to affect the toxicity of the oil they are applied to. As a result, the use of oil dispersants across the globe has been historically controversial, especially in light of large scale, high profile spill events from the Torrey Canyon spill in 1967 to the Deepwater Horizon incident in 2010. In the case of the Torrey Canyon disaster, this was due to the use of first-generation oil dispersants, primarily consisting of industrial detergents. In recent years, developments in the production of dispersants has resulted in the far less toxic second and third generation oil dispersants which generally are comprised of a surfactant and a solvent. The solvent facilitates the transport of the surfactant through the oil layer to the oil-water interface where it reduces the interfacial tension between the oil and water, encouraging the oil to disperse into smaller droplets. Whilst these smaller droplets are more bioavailable to marine organisms (Rico-martínez et al., 2013), the chemical dispersion results in faster dilution and biodegradation of the oil.

Despite the evolution of oil dispersants, there is still a need to ensure the most environmentally acceptable products (e.g. lowest toxicity) are approved for use. It is also worth noting that dispersants are not the only oil spill treatment products on the market; demulsifiers, surface cleaners, bioremediation products, sorbents and degreasers are also options and appropriate under specific scenarios. The choice of product depends on factors such as oil type, sea conditions and the surface that the oil has come into contact with (be it water, rocky shore or other) and therefore regimes for assessing the use of oil spill treatment products also need to take into account these scenarios. In the UK, in order to have the most efficient products with the lowest possible toxicity, products must go through a pre-approval system before they can be marketed and, ultimately, stockpiled for mobilisation in the event of an oil spill. Current

experimental protocols for the toxicity assessment of oil spill treatment products in the UK have been established since the 1970s (Blackman et al., 1978) and have not been updated since 1996 (Kirby et al., 1996). To address health and safety, cost and scientific robustness issues, the UK approach for dispersant testing and approval has been reviewed and updated for implementation during 2020.

To make an informed decision on the most appropriate methods to use for future toxicity assessments of oil spill treatment products for use in UK waters, a number of current approaches, including those used by regulatory authorities in other jurisdictions, were reviewed to determine those most likely to provide reproducible, reliable, interpretable and cost-effective results.

METHODS

Choice of species

Historically many toxicity assessments of oils and dispersants have been conducted as part of research in relation to large, high profile oil spill events and are usually unrelated (i.e. different test items and species) and experimentally diverse (i.e. different methods of exposure or test item preparation) creating a lot of data that is hard to interpret and therefore of limited value for risk assessment process (Lewis and Pryor, 2013; Redman and Parkerton, 2015). Studies have shown that closely related species generally share similar toxicity values (Hansen et al., 2014; Wilson et al., 1973) and in others there were no apparent taxonomic trends (Barron et al., 2013; George and Clark, 2000) depending on whether oil was tested in combination with dispersants or oil and oil dispersants were tested independently. Given this lack of consistently, an assessment of 17 studies evaluating the impact of oil and oil dispersants was conducted, in an attempt to highlight any taxonomic trends (Figure 1). Here, the concentration of test item to cause a response in 50% of the test species population (EC_{50} value) ranged from 0.03 mg/L to

788 mg/L and generally EC_{50}'s were found to be above 10 mg/L for dispersants, below 10 mg/L of dispersant where oil is assessed in combination with the dispersant or below 10 mg/L of oil where oil is assessed without a dispersant (it should be noted that in some papers oil concentration is as mg total oil/ L and in others it is Total Petroleum Hydrocarbons/ L) (Adams et al., 2014; Almeda et al., 2014; Aurand and Coelho, 2005; Clark et al., 2001; Coelho and Aurand, 1996; Cotou et al., 2001; Edwards et al., 2003; Fuller et al., 2004; Hemmer et al., 2011; Mitchell and Holdway, 2000; Rhoton, 1999; Scarlett et al., 2005; Singer et al., 1996, 1993, 1991, 1990; U.S. Environmental Protection Agency, 2010).

The greatest amount of data exists for fish and crustaceans, which is likely as a result of Mysids and Menidia having been used as standard test species in oil toxicology for decades as the recommended species for dispersant tests in the US National Contingency Plan (NCP) (Hemmer et al., 2011). However, the wide range of EC_{50}'s shown in Figure 1, even when comparing results for the same product, is probably due to a lack of standardisation in the methods used i.e. different chemical preparation methods, tests not run to standard guidelines, standard guidelines adapted in different ways, different life stages from embryo to adult, organisms from different classes within taxonomic groups, various end points e.g. mortality, growth, development, oxygen consumption, response to stimulus, different types of oil, different dispersants and differing durations, so comparisons are difficult to make. This means that the most reliable methods and species to use in oil dispersant risk assessment was not clear.

Despite the conflicting data on the scale of sensitivity difference across taxa, it is only logical that test organism selection should represent those most likely to be indigenous to the environments in which oil spills and the use of oil dispersants are most likely (Posthuma and de Zwart, 2012; Suter et al., 2002; Wu, 1981). Generally dispersants are only found in the top 5-10 m of water (Singer et al., 1991) meaning that among the biological components of marine ecosystems, planktonic organisms are those that are particularly vulnerable to crude oil

pollution (Almeda et al., 2014; Graham et al., 2010; Walsh, 1978). Perhaps even more pertinent however, is the use of standard species and guidelines to ensure ease of interpretation and reproducibility (The National Academies of Sciences Engineering and Medicine, 2019). It is impossible for laboratory tests to reflect the exact dilution conditions and mixing energy of the sea, as a result it is important that laboratory testing methodologies

do not introduce arbitrary restrictions that relate to effects that are an artefact of the laboratory conditions as opposed to environmental conditions (Lunel, 2001). Therefore, it was concluded that any test species selected should represent where oil and dispersants are likely to be used, but also have standard testing guidelines associated with them in order to meet the intended objective of a reliable, reproduceable and cost-effective test.

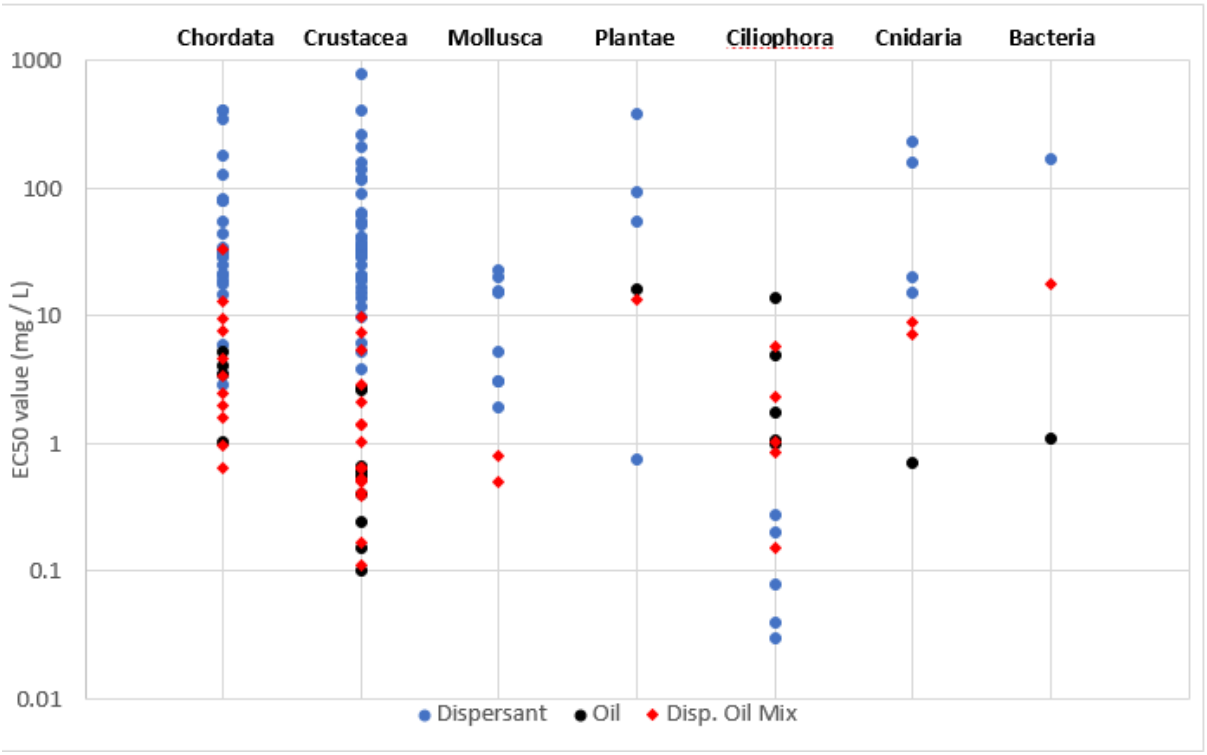

Figure 1. The toxicity of oil, dispersants and oil and dispersant mixes to different taxonomic groups, generated from 17 peer reviewed studies, totalling 160 independent EC50 values.

In a regulatory context, the most commonly used organisms for toxicity assessments are fish (various), crustaceans, molluscs (oysters or mussels) and algae (e.g. *Skeletonema* sp). Although a plethora of information exists for the toxicity of oil and dispersants to fish, the

strong ethical argument for avoiding the use of vertebrates in toxicity testing and, given that often invertebrates are shown to be more sensitive, meant that the use of fish can be avoided whilst still selecting suitable species. As a result, the use of fish was excluded from the proposed assessment scheme. That said, it is advantageous to cover different trophic levels in toxicity-based assessment schemes as using only one trophic level could reduce the power of the approach and have knock-on effects for the whole food web (Biddinger and Gloss, 1984).

The standard acute toxicity tests most commonly used for pelagic marine invertebrates in the UK are the Pacific oyster (*Crassostrea gigas*) and mussel (*Mytilus edulis or Mytilus galloprovincialis)* embryo-larval development test and the marine copepod (*Acartia tonsa, Tisbe battagliai* or *Nitocra spinipes*) acute lethal toxicity test. The use of bivalves in toxicity testing can be limited by reproductive seasonality and maintenance of viable larvae (Bellas and Paredes, 2011; Mcfadzen, 1992). Developmental abnormalities of fertilised oyster eggs can be as high as 50 % in controls (Thain, 1991; Utting and Spencer, 1991). Such high control abnormalities are likely to be congenitally derived (Bayne, 1972), and there is no means of detecting this prior to the onset of the bioassay. These issues can result in unreliable data and rapidly increasing costs due to the requirement for repeat testing. Given the need for cost effective tests throughout the year, they were deemed not suitable for use in this testing scheme. The copepod *Tisbe battagliai* is routinely used for chemical risk assessment in Northern Europe and its relevance to chemical pollution in the water column has been demonstrated by a number of research papers (Thomas et al., 1999) and therefore is a sensible choice for toxicity testing. As copepods are primary consumers, feeding on algae, they form a meaningful link between the algae and higher trophic levels. Impacts at this level of the food web can have important consequences for the higher, more 'visible' species such as fish (copepods are an important part of the diet of larval fish) and birds.

Toxicity data for the effects of oil dispersants on algae is more limited, especially for standardised test methods. However, the literature suggests that, overall, algae generally has a similar or lower sensitivity than fish and pelagic crustacea (Barron et al., 2013; Hansen et al., 2014; Lewis and Pryor, 2013). That said, algae sensitivity is dependent on dispersant type and it has been found to be amongst the most sensitive species to Corexit 9500 and 9527 (Bejarano, 2018; Echols et al., 2019). It is important to include algae in an oil spill treatment product toxicity testing regime as it forms the base of the ecosystem, has a widespread distribution and has the potential to impose secondary effects on higher trophic levels (Brooks et al., 2008). Of the EU and EFTA (European Free Trade Association) countries that have oil spill treatment product approval schemes the most commonly used species is the planktonic algae *Skeletonema* sp. (European Maritime Safety Agency, 2009). Given the availability of standard guidelines and their relevance to UK waters, whilst posing no ethical questions or seasonality, *Tisbe battagliai* and *Skeletonema* sp. were selected as the most appropriate species to proceed to robustness testing.

Inclusion of oil

The selection of standard methods provides criteria to follow in test set up. However, when considering combined oil and dispersant testing, the preparation method for the test solutions can significantly affect toxicant availability/exposure and therefore needs to be carefully considered (Redman and Parkerton, 2015). When including oil in the toxicity assessment copious research has already been conducted by the Chemical Response to Oil Spills Environmental Research Forum (CROSERF) who used evidence from interlaboratory comparisons in order to recommend a standardised design that would facilitate consistent oil toxicity test results. The most important aspects to consider have been identified as the following: preparation of test solutions, analytical confirmation of test concentrations and

method of exposure (Singer et al., 2000). When assessing oil alone a Water Accommodated Fraction (WAF) is used, in order to avoid emulsification, a lower energy WAF than is recommended for normal toxicity testing is required, low enough to produce no vortex. This also makes the procedure more replicable, and applicable to a wide range of oils and other emulsifying substances.

When assessing the dispersant in combination with the oil a Chemically Enhanced Water Accommodated Fraction (CEWAF) is used. When preparing a CEWAF the mixing energy needs to be sufficient to create a vortex in order for appropriate interactions of the dispersant and the oil. In addition to this, the method of dispersant delivery and order of oil/dispersant addition, mixing initiation and settling duration is very important in a CEWAF and different to that of the oil WAF (Singer et al., 2000).

Where oil dispersants are assessed alone, there is no specific guidance on preparation of the test media for toxicity testing. For soluble oil dispersants standard dilution in sea water from a stock solution is appropriate providing the resulting concentrations are not so high as to be above the solubility limit of the test substance. Where resulting concentrations in the test media exceed the solubility limit, so that undissolved test substance is apparent, use of a WAF dosing method as suggested in ISO 14442 (1999) would be suitable. When dealing with oil spill treatment products other than dispersants, preparation will be dictated by the process of application described by the supplier and may require amendments to standard methods.

In the case of test media preparation for oil toxicity tests, to simulate the multicomponent dissolution behaviour of the hydrocarbon constituents that vary widely in physical chemical properties, a variable oil loading approach is preferred (Girling et al., 1994; Singer et al., 2000). This avoids the complications that arise when using variable (serial) dilution test protocols as discussed in The Use of Dispersants in Marine Oil Spill Response (The National Academies of Sciences, Engineering and Medicine, 2019).).

Much of the toxicity associated with the use of oil spill treatment products is due to the dispersed oil rather than the product itself and oil of different concentrations or types (e.g. crude or light fuel oil) can have widely differing volatile components (Singer et al., 2000; The National Academies of Sciences Engineering and Medicine, 2019). Similarly, the increase in toxicity of dispersed oil (chemically or mechanically) compared to undispersed oil can be attributed to particle size in the dispersion (Bobra et al., 1989) and aromatic hydrocarbon content (Anderson et al., 1987), but also to a number of interacting chemical, physical and physiological factors (The National Academies of Sciences Engineering and Medicine, 2019). Therefore, with the need to prepare oil differently to an oil and dispersant mixture, or dispersant only solution, along with the potential for differing dissolved oil and microdroplet exposures, in order to ensure reproducibility, this introduces uncertainty as to the suitability of including oil at all in oil spill treatment product toxicity assessments.

The aim of dispersants is to stop oil reaching sensitive habitats and, therefore, approval schemes normally include an assessment of efficiency. In general, the more efficient a product is, the more toxic the dispersed oil and dispersant mix will be and therefore including dispersed oil in the consideration could lead to the rejection of the most efficient products (Interspill, 2012). Although desirable to mimic real environmental conditions, as much as possible, with oil and dispersant testing there are so many variables that ensuring the test is interpretable and repeatable is recognised as being more important in order to generate meaningful data (The National Academies of Sciences Engineering and Medicine, 2019). When looking at regulators in jurisdictions other than the UK only two out of 12 countries appear to include oil in their approval schemes and, where the information is available, the others set threshold limits for the toxicity of the dispersant alone (with the exception of the French regulatory authorities who set a criterion that toxicity must be 10x less than a reference chemical) (European Maritime Safety Agency, 2009).

Testing of dispersants

In order to assess the selected test methods and affirm the decision to exclude oil from the testing regime, eight dispersants, commonly stockpiled in the UK, were tested independently and two dispersants in combination with a representative crude oil (Kuwait). Core test species were the copepod, *Tisbe battagliai*, and the algae, *Skeletonema* sp. Dispersant only tests were conducted according to ISO 10253 (2016) and ISO 14669 (1999). Dispersant and oil tests were conducted at a 1:10 dispersant:oil ratio with *Tisbe battagliai* only, according to ISO 14669 (1999) with test item preparation conducted according to Singer et al., (2000). All tests were conducted under a static regime, with exposure durations of 48 hours and 72 hours respectively for *Tisbe battagliai* and *Skeletonema* sp. tests. Results are displayed as the nominal concentration of dispersant to cause mortality in 50% of the test organisms (LC_{50}) in the case of *Tisbe battaliai* or 50% growth inhibition (IC_{50}) in the case of *Skeletonema* sp. With each dispersant test a reference test using 3,5-dichlorophenol was conducted using the same algae or copepod culture to check its sensitivity and ensure the culture was not compromised but acting normally. All *Tisbe battagliai* reference tests had nominal $EC_{50's}$ between 0.71 and 1.71 mg/L 3,5-dichlorophenol. All *Skeletonema* sp. reference tests had nominal $EC_{50's}$ between 0.59 and 2.56 mg/L 3,5-dichlorophenol. These results are within the in-house control chart limits for these tests.

RESULTS/DISCUSSION

Inclusion of oil

When comparing the toxicity of oil alone to a 1:10 dispersant:oil mix using CROSERF WAF and CEWAF methods, no toxicity to *Tisbe battagliai* was seen in the oil only test due to the low stirring velocity of the WAF (i.e. stirring with low/no vortex to avoid

emulsion forming (~100RPM)). It is likely that CEWAF test media preparation results in faster dissolution and higher dissolved oil exposures than the WAF test media preparation (The National Academies of Sciences Engineering and Medicine, 2019). When mechanically dispersed, oil toxicity was only observed when the stirring speed of the WAF was relatively high (350RPM). This presents problems for an assessment regime comparing oil and dispersant to oil only toxicity; if the toxicity of the dispersant in combination with the oil must be of a lower toxicity than the oil alone, all dispersants would fail if CROSERF methods are followed and compared using the approach illustrated in Figure 2.

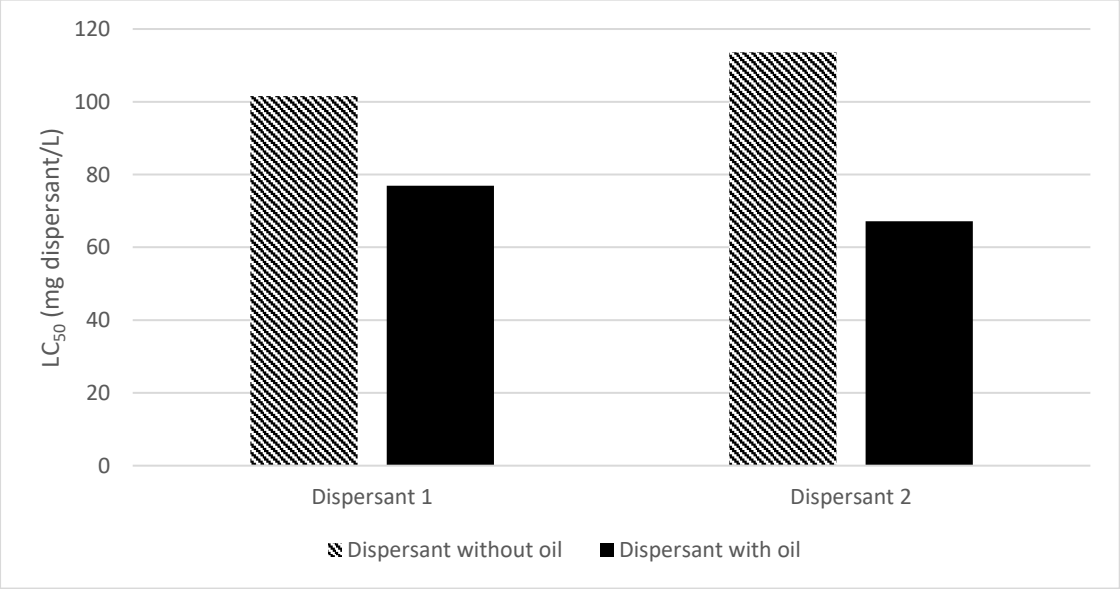

Figure 2 LC_{50} values for the toxicity of two oil dispersants to Tisbe battagliai both with and without Kuwait crude oil. LC_{50} indicates concentration of dispersant in mg/L when applied at a ratio of 1:10 Dispersant:Oil

Figure 2 shows that tests with oil in combination with a dispersant were approximately 30% more toxic to *Tisbe battagliai* than the dispersant alone. This is consistent with previous studies and suggests that the increased toxicity seen in the dispersant and oil treatments is due to the increase in biological availability of oil constituents (Adams et al., 2014; Fisher and Foss, 1993; Fuller et al., 2004; Lindén, 1975). Generally, increases in chemically dispersed oil toxicity are directly correlated to the effectiveness of the dispersant, due to an increase in solubilised toxic hydrocarbons (Fisher and Foss, 1993). As a result, the observed increases in

toxicity in an oil and dispersant mix may not be a representation of the combined toxicity of the oil and dispersant but may be an indication of the effectiveness of the dispersant (Adams et al., 2014).

It has also been demonstrated that the use of different oils (i.e. source, age, composition) has produced results that are impossible to compare to one another due to differences in the inherent toxicities of different oils (Holder et al., 2015; The National Academies of Sciences Engineering and Medicine, 2019), so the use of one reference oil may produce different results to the use of another. This is demonstrated by the order of magnitude difference in observed toxicity of the dispersant Corexit 9500 and three different oils as reported by Ramachandran et al. (2004). The composition of an oil is directly related to its toxicity, particularly Polycyclic Aromatic Hydrocarbon (PAH) levels, and as such unless the same oil (i.e. source, age, composition) is used, it is impossible to compare results between tests due to differences in the dissolved hydrocarbons present. The test system can also have a significant effect on the toxicity of the oil with open systems resulting in the loss of potentially toxic volatile components of the oil and variable (serial) dilutions resulting in differences in the presence of microdroplets and thus higher dissolved fractions at lower dilutions than variable loading test media preparations (The National Academies of Sciences Engineering and Medicine, 2019). As a result, the use of oil is inappropriate for an easily interpretable and reliable oil spill treatment product toxicity screening test.

Reproducibility and threshold limit

When assessing the toxicity of eight dispersants to *Skeletonema* sp. (**Error! Reference source not found.**), all but one had 50% inhibition concentrations above 10 mg/L. In order to assess the reproducibility of the method three dispersants were chosen, for which the *Skeletonema* sp. assay was conducted, independently, multiple times (dispersants 1, 2 and 3).

The means of the independently conducted assays were compared using a Wilcoxon Mann Whitney non-parametric (WMW) test. This test is suitable for the analysis of this type of data because it is particularly robust for small datasets and does not assume normality or heterogeneity of variance.

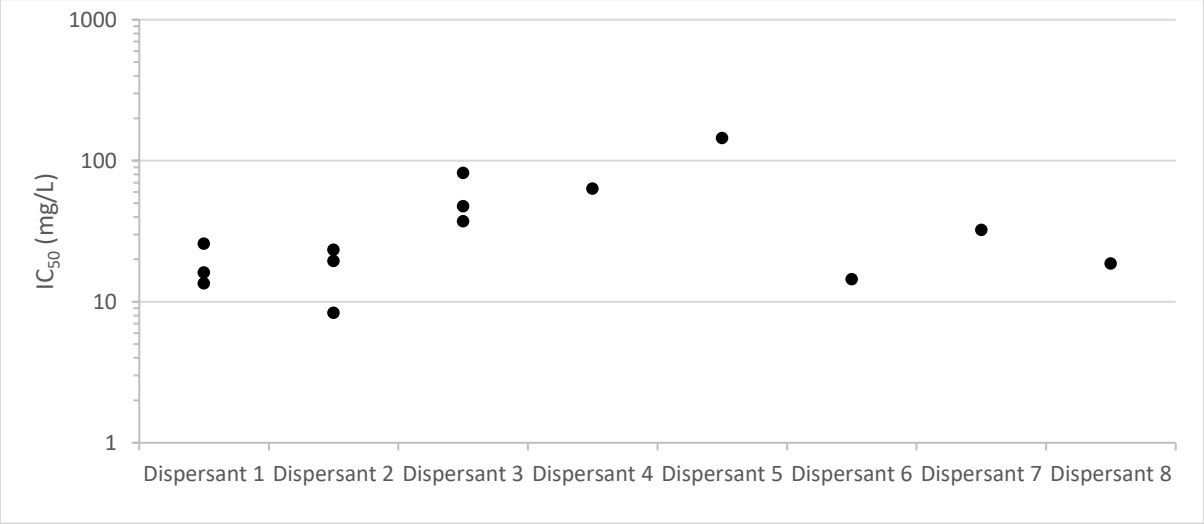

Figure 3. IC_{50} values for the toxicity of eight dispersants to Skeletenema sp.

The WMW test found that the difference between the means were not statistically significant (

Table 1). Therefore, in all cases we cannot reject H0: meaning that, the IC50's generated for *Skeletonema* sp. assays run independently with the same dispersant were consistent. When conducted in this way the *Skeletonema* sp. assay is a reproducible method, suitable for assessing the toxicity of oil dispersants.

The toxicity of the same dispersants (with the exception of dispersant eight for which data is not available) to *Tisbe battagliai* on average was found to be lower than to *Skeletonema* sp. **Error! Reference source not found.** shows that all *T. battagliai* LC_{50}s were above 10 mg/L. Oil dispersant regulators in Norway and New Zealand use a pass/fail threshold of 10 mg/L regardless of the test species (Interspill, 2012). When considering where to best set a threshold concentration based on the data gathered here, as well as that in the literature, a threshold of 10 mg/L would ensure that a wide range of efficient products would be available for use whilst also protecting at-risk species since dispersant concentrations in the field are expected to me orders of magnitude lower.

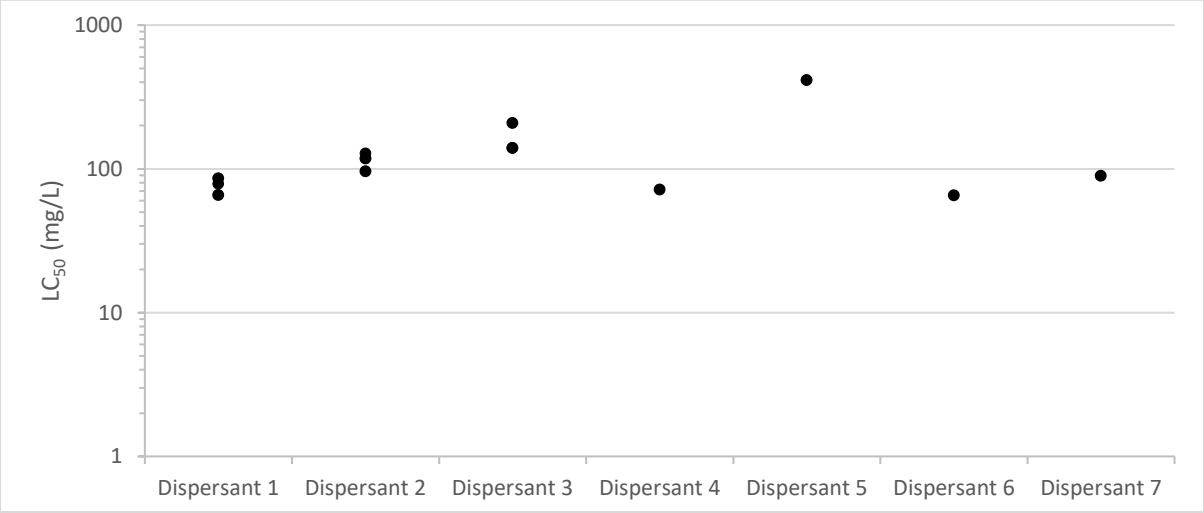

Figure 4 LC_{50} values for the toxicity of eight dispersants to Tisbe battagliai

In both the *T. battagliai* and *Skeletonema* sp. bioassays Dispersant 6 was, on average, shown to be the most toxic and Dispersant 5 the least toxic, showing similar sensitivities in the two species. As for *Skeletonema* sp., in order to assess the reproducibility of the *T. battagliai*

method, Dispersants 1, 2 and 3 were tested multiple times. Again, to compare the means of the tests run independently we used the four replicates (used to calculate each mean) in the Wilcoxon Mann Whitney non-parametric (WMW) test.

The WMW test found that the difference between the mean LC_{50}'s were not statistically significant (

Table 1). Therefore, in all cases we cannot reject H0: indicating that, statistically, the LC_{50}'s were the same for each dispersant. This indicates that the *T. battagliai* assay methods employed here were consistently reproducible in independent tests and are therefore suitable in assessing the toxicity of oil dispersants.

Table 1 Results of Wilcoxon Mann Whitney non-parametric statistical analysis conducted for T. battagliai and Skeletonema sp. *with Dispersants 1, 2 and 3*

Species	Dispersant	Test number	W index	p-value	Significance
Skeletonema sp.	1	Test 1 vs Test 2	9	0.1	Non significant
Skeletonema sp.	1	Test 1 vs Test 3	9	0.1	Non significant
Skeletonema sp.	1	Test 2 vs Test 3	2	0.4	Non significant
Skeletonema sp.	2	Test 1 vs Test 2	9	0.1	Non significant
Skeletonema sp.	2	Test 1 vs Test 3	8	0.2	Non significant
Skeletonema sp.	2	Test 2 vs Test 3	0	0.1	Non significant
Skeletonema sp.	3	Test 1 vs Test 2	9	0.1	Non significant
Skeletonema sp.	3	Test 1 vs Test 3	9	0.1	Non significant
Skeletonema sp.	3	Test 2 vs Test 3	9	0.1	Non significant
T. battagliai	1	Test 1 vs Test 2	13	0.1635	Non significant
T. battagliai	1	Test 1 vs Test 3	14	0.01016	Non significant
T. battagliai	1	Test 2 vs Test 3	7.5	1.0	Non significant
T. battagliai	2	Test 1 vs Test 2	4	0.3094	Non significant
T. battagliai	2	Test 1 vs Test 3	10.5	0.559	Non significant
T. battagliai	2	Test 2 vs Test 3	14.5	0.08143	Non significant
T. battagliai	3	Test 1 vs Test 2	12	0.04975[1]	Non significant
T. battagliai	3	Test 1 vs Test 3	10	0.2076	Non significant
T. battagliai	3	Test 2 vs Test 3	9	0.8824	Non significant

[1]The difference of the means of these two tests was checked using the alternative non-parametric test of Kolmogorov-Smirnov which confirmed that we could not reject H0 $p<0.06$, and that the test results were not significantly different

CONCLUSIONS

The aim of dispersants is to reduce the amount of oil reaching sensitive habitats and to promote biodegradation. Product approval schemes generally aim to select the appropriately efficient products, but the nature of dispersants mean that the more efficient a product is the more toxic the dispersed oil and dispersant mixture will be when compared to dispersant only toxicity. Therefore, including the toxicity of dispersed oil in a scheme could lead to the rejection of the most efficient products. Similarly, the decision to use an oil spill treatment product should not be based solely on its reported aquatic toxicity, but as a wider risk assessment and therefore highlights the need, in a regulatory context, to only exclude the most toxic of the most efficient products. It is impossible for laboratory tests to reflect the exact dilution conditions and mixing energy of the sea; therefore, it is important that laboratory testing methodologies do not introduce arbitrary restrictions that relate to effects that are an artefact of the laboratory conditions as opposed to environmental conditions. This study has

demonstrated that two environmentally relevant species, *Skeletonema* sp. and *Tisbe battagliai,* that can be easily held/cultured in the laboratory with no seasonality and with cost-effective standardised methods associated with them, can produce reproduceable and interpretable results to ensure the most toxic products are identified.

Standardisation of oil spill treatment product testing will not only be safer and more cost-effective but will promote more effective cross-institute comparisons of toxicity data and will enable further harmonisation of approaches in the future. The redevelopment of the UK guideline to use standardised testing and the selection of appropriate, environmentally relevant test organisms will increase the quality and reliability of data used to underpin the UK oil spill treatment testing and approval scheme. This in turn will lead to more appropriate, well informed decisions being made that are backed with robust, reliable scientific data.

The adoption of this approach will enable an approved list of products for use in UK waters to be maintained. However, the decision for dispersant use in any given scenario will still need to be underpinned through expert advice applying a risk assessment approach taking account a range of incident-specific physical and environmental sensitivity information.

REFERENCES

Adams, J.A., Sweezey, M.S., Hodson, P.V.H., 2014. Oil and oil dispersant do not cause synergistic toxicity to fish embryos. Environ. Toxicol. Chem. 33, 107–114. https://doi.org/10.1002/etc.2397

Almeda, R., Hyatt, C., Buskey, E.J., 2014. Toxicity of dispersant Corexit 9500A and crude oil to marine microzooplankton. Ecotoxicol. Environ. Saf. 106, 76–85. https://doi.org/10.1016/j.ecoenv.2014.04.028

Anderson, J.W., Riley, R., Kiesser, S., Gurtisen, J., 1987. Toxicity of dispersed and undispersed Prudhoe Bay crude oil fractions to shrimp and fish, in: International Oil

Spill Conference Proceedings. pp. 235–240.

Aurand, D., Coelho, G. (Editors), 2005. Cooperative Aquatic Toxicity Testing of Dispersed Oil and the "Chemical Response to Oil Spills: Ecological Effects Research Forum (CROSERF)." Ecosyst. Manag. Assoc. Inc. Lusby, MD. Tech. Rep. 07-03, 105 pages + Append.

Barron, M.G., Hemmer, M.J., Jackson, C.R., 2013. Development of aquatic toxicity benchmarks for oil products using species sensitivity distributions. Integr. Environ. Assess. Manag. 9, 610–615. https://doi.org/10.1002/ieam.1420

Bayne, B.L., 1972. Some Effects of Stress in the Adult on the Larval Development of Mytilus edulis. Nature 237, 459.

Bejarano, A.C., 2018. Critical Review and Analysis of Aquatic Toxicity Data on Oil Spill Dispersants. Environ. Toxicol. Chem. 37, 2989–3001. https://doi.org/10.1002/etc.4254

Bellas, J., Paredes, E., 2011. Cryobiology Advances in the cryopreservation of sea-urchin embryos : Potential application in marine water quality assessment q. Cryobiology 62, 174–180. https://doi.org/10.1016/j.cryobiol.2011.02.005

Biddinger, G.R., Gloss, S.P., 1984. The importance of trophic transfer in the bioaccumulation. Residue Rev. 91.

Blackman, R.A.A., Franklin, F.L., Norton, M.G., Wilson, K.W., 1978. New Procedures for the Toxicity Testing of Oil Slick Dispersants in the United Kingdom. Mar. Pollut. Bull. 9, 234–238.

Bobra, A.M., Shui, W.Y., Mackay, D., Goodman, R.H., 1989. Acute toxicity of dispersed fresh and weathered crude oil and dispersant to daphnia magna. Chemosphere 19, 1199–1222.

Brooks, S.J., Bolam, T., Tolhurst, L., Bassett, J., La, J., Waldock, M., Barry, J., Thomas, K. V, 2008. Dissolved organic carbon reduces the toxicity of copper to germlings of the

macroalgae , Fucus vesiculosus $ 70, 88–98. https://doi.org/10.1016/j.ecoenv.2007.04.007

Clark, J.R., Bragin, G.E., Febbo, E.J., Letinski, D.J., 2001. Toxicity of physically and chemically dispersed oils under continuous and environmentally realistic exposure conditions: Applicability to dispersant use decisions in spill response planning, in: International Oil Spill Conference. Global Strategies for Prevention, Preparedness, Response, and Restoration. API Publication No. 4686B (Same Number Used for the 1999 Proceedings). American Petroleum Institute, Washington, D.C. pp. 1249–1255.

Coelho, G.M., Aurand, D. V., 1996. Proceeding of the Fifth Meeting of the Chemical Response to Oil Spills: Ecological Effects Research Forum., in: Corpus Christi, Texas, USA. Ecosystem Management & Associates, Purcellville, VA, Report 96-03, 60 P. pp. 18–19.

Cotou, E., Castritsi-Catharios, I., Moraitou-Apostolopoulou, M., 2001. Surfactant-based oil dispersant toxicity to developing nauplii of Artemia : effects on ATPase enzymatic system. Chemosphere 42, 959–964.

Echols, B., Langdon, C., Stubblefield, W., Rand, G., Gardinali, P., 2019. A Comparative Assessment of the Aquatic Toxicity of Corexit 9500 to Marine Organisms. Arch. Environ. Contam. Toxicol. 77, 40–50.

Edwards, K.R., Lepo, J.E., Lewis, M.A., 2003. Toxicity comparison of biosurfactants and synthetic surfactants used in oil spill remediation to two estuarine species. Mar. Pollut. Bull. 46, 1309–1316. https://doi.org/10.1016/S0025-326X(03)00238-8

European Maritime Safety Agency, 2009. Manual on the Applicability of Oil Spill Dispersants - Version 2.

Fisher, W.S., Foss, S.S., 1993. A Simple Test for Toxicity of Number 2 Fuel Oil and Oil Dispersants to Embryos of Grass Shrimp , Palaemonetes Pugio. Mar. Pollut. Bull. 26,

385–391.

Fuller, C., Bonner, J., Page, C., Ernest, A., McDonald, T., McDonald, S., 2004. Comparative toxicity of oil, dispersant, and oil plus dispersnt to several marine species. Environ. Toxicol. Chem. 23, 2941–2949.

George, A., Clark, J.R., 2000. Aquatic toxicity of two Corexit dispersants. Chemosphere 40, 897–906.

Girling, A.E., Girling, A., Consultant, E., Whale, G., Health, S., 1994. A guideline supplement for determining the aquatic toxicity of poorly water-soluble complex mixtures using water- accommodated fractions. https://doi.org/10.1016/0045-6535(94)90065-5

Graham, W.M., Condon, R.H., Carmichael, R.H., Ambra, I.D., Patterson, H.K., Linn, L.J., 2010. Oil carbon entered the coastal planktonic food web during the Deepwater Horizon oil spill. https://doi.org/10.1088/1748-9326/5/4/045301

Hansen, B.H., Altin, D., Bonaunet, K., Overjordet, I.B., 2014. Acute Toxicity of Eight Oil Spill Response Chemicals to Temperate, Boreal and Arctic Species. J. Toxicol. nd Environemtnal Haleath, Part A 77, 495–505.

Hemmer, M.J., Barron, M.G., Greene, R.M., 2011. Comparative toxicity of eight oil dispersants, Louisiana sweet crude oil (LSC), and chemically dispersed LSC to two aquatic test species. Environ. Toxicol. Chem. 30, 2244–2252. https://doi.org/10.1002/etc.619

Holder, E.L., Conmy, R.N., Venosa, A.D., 2015. Comparative Laboratory-Scale Testing of Dispersant Effectiveness of 23 Crude Oils Using Four Different Testing Protocols * 628–639.

Interspill, 2012. Oil spill dispersants : Developments in regulations and industry perspectives, in: A Background Discussion Paper for a Dispersant Session at Interspill 2012. pp. 1–18.

ISO 10253, 2016. Water quality - Marine algal growth inhibition test with Skeletonema sp . and Phaeodactylum tricornitum.

ISO 14442, 1999. Water quality — Guidelines for algal growth inhibition tests with poorly soluble materials, volatile compounds, metals and waste water.

ISO 14669, 1999. Water quality - Determination of acute lethal toxicity to marine copepods (Copepoda, Crustacea).

ITOPF, 2019. Oil Tanker Spill Statistics 2018.

Kirby, M.F., Matthiessen, P., Rycroft, R.J., 1996. Procedures for the approval of oil spill treatment products, Fisheries Research Technical Report.

Lewis, M., Pryor, R., 2013. Toxicities of oils , dispersants and dispersed oils to algae and aquatic plants : Review and database value to resource sustainability. Environ. Pollut. 180, 345–367. https://doi.org/10.1016/j.envpol.2013.05.001

Lindén, O., 1975. Acute Effects of Oil and Qil / Dispersant Mixture on Larvae of Baltic Herring. Ambio 4, 130–133.

Lunel, T., 2001. Dispersant Pre-Approvals-Best Practice, in: International Oil Spill Conference Proceedings: March 2001. pp. 441–444.

Mcfadzen, I.R.B., 1992. Growth and survival of cryopreserved oyster and clam larvae along a pollution gradient in the German Bight 91, 215–220.

Mitchell, F.M., Holdway, D.A., 2000. The acute and chronic toxicity of the dispersants Corexit 9527 and 9500, water accommodated fraction (WAF) of crude oil, and dispersant enhanced WAF (DEWAF) to Hydra viridissima (green hydra). Water Res. 34, 343–348. https://doi.org/10.1016/S0043-1354(99)00144-X

Posthuma, L., de Zwart, D., 2012. Predicted mixture toxic pressure relates to observed fraction of benthic macrofauna speces impacted by contaminant mixtures. Environ. Toxicol. Chem. 31, 2175–2188. https://doi.org/10.1002/etc.1923

Ramachandran, S.D., Hodson, P. V, Khan, C.W., Lee, K., 2004. Oil dispersant increases PAH uptake by fish exposed to crude oil. Ecotoxicol. Environ. Saftey 59, 300–308. https://doi.org/10.1016/j.ecoenv.2003.08.018

Redman, A.D., Parkerton, T.F., 2015. Guidance for improving comparability and relevance of oil toxicity tests. Mar. Pollut. Bull. 98, 156–170. https://doi.org/10.1016/j.marpolbul.2015.06.053

Rhoton, S.L., 1999. Acute toxicity of the oil dispersant corexit 9500, and fresh and weathered Alaska North Slope crude oil to the Alaskan tanner crab (C. Bairdi), two standard test species, and V. fischeri (MICROTOX Assay). University of Alaska, Fairbanks.

Rico-martínez, R., Snell, T.W., Shearer, T.L., 2013. Synergistic toxicity of Macondo crude oil and dispersant Corexit 9500A Ò to the Brachionus plicatilis species complex (Rotifera). Environ. Pollut. 173, 5–10. https://doi.org/10.1016/j.envpol.2012.09.024

Scarlett, A., Galloway, T.S., Canty, M.C., Smith, E.L., Nilsson, J., Rowland, S.. J., 2005. Comparative toxicity of two oil dispersants, Superdispersant-25 and Corexit 9527, to a range of coastal species. Environ. Toxicol. Chem. 24, 1219–1227.

Singer, M.M., Aurand, D., Bragin, G.E., Clark, J.R., Coelho, G.M., Sowby, M.L., 2000. Standardization of the Preparation and Quantitation of Water-accommodated Fractions of Petroleum for Toxicity Testing. Mar. Pollut. Bull. 40, 1007–1016.

Singer, M.M., George, S., Benner, D., Jacobson, S., Tjeerdema, R.S., Sowby, M.L., 1993. Comparative toxicity of two oil dispersants to the early life stages of two marine species. Environ. Toxicol. Chem. 12, 1855–1863.

Singer, M.M., George, S., Jacobson, S., Lee, I., Weetman, L.L., Tjeerdema, R.S., Sowby, M.L., 1996. Comparison of Acute Aquatic Effects of the Oil Dispersant Corexit 9500 with Those of Other Corexit Series Dispersants. Ecotoxicol. Environ. Saftey 35, 183–189.

Singer, M.M., Smalheer, D.L., Tjeerdema, R.S., 1991. Effects of spiked exposure to an oil dispersant on the early life stages of four marine species. Environ. Toxicol. Chem. 10, 1367–1374.

Singer, M.M., Smalheer, D.L., Tjeerdema, R.S., Martin, M., 1990. Toxicity of an oil dispersant to the early life stages of four California marine species. Environ. Toxicol. Chem. 9, 1387–1395.

Suter, G., Traas, T., Posthuma, L., 2002. Issues and practices in the derivation and use of species sensitivity distributions., in: Posthuma, L., Suter, G., Traas, T. (Eds.), Species Sensitivity Distributions in Ecotoxicology. CRC Press, pp. 437–474.

Thain, J.E., 1991. Biological effects of contaminants: oyster (Crassostrea gigas) embryo bioassay. ICES Report. Tech. Mar. Environ. Sci.

The National Academies of Sciences Engineering and Medicine, 2019. The Use of Dispersants in Marine Oil Spill Response. Washington, DC: The National Academies Press.

Thomas, K.V., Thain, J.E., Waldock, M.J., 1999. Identification of toxic substances in UK estuaries. Environ. Toxicol. Chem. 18, 401–411.

U.S. Environmental Protection Agency, 2010. National Oil and Hazardous Substances Pollution Contingency Plan 40 CFR 300. U.S. Government Printing Office, Washington, DC.

Utting, S.D., Spencer, B.E., 1991. The hatchery culture of bivalve mollusc larvae and juveniles.

Walsh, G.E., 1978. Toxic Effects of Pollutants on Plankton, in: Butler, G.C. (Ed.), Principles of Ecotoxicology. John Wiley and Sons, Inc, New York, pp. 257–274.

Wilson, K.W., Cowell, E.B., Beynon, L.R., 1973. The Toxicity Testing of Oils and Dispersants: A European View. Int. Oil Spill Conf. Proc. March 1973 1973, 255–261.

Wu, R.S.S., 1981. Differences in the toxicities of an oil dispersant and a surface active agent to some marine animals, and their implications in the choice of species in toxicity testing. Mar. Environ. Res. 5, 157–163.