Umwelt Bundes Amt
For our Environment

Question-and-Answer Document by the German Federal Environment Agency on the new OECD Guideline 209 “Activated Sludge, Respiration Inhibition Test (Carbon and Ammonium Oxidation) (22 July 2010)”

Version
02.03.2012

Since the new version of OECD 209 entered into force, the German Federal Environment Agency (Umweltbundesamt, UBA) has received various enquiries from applicants and test institutes concerning the interpretation of the Guideline and the resulting performance of the test. Some passages in the new Guideline need to be revised since they are redundant, not scientifically sound or insufficiently explained. During the period for comment on the draft OECD guideline, the Federal Environment Agency submitted critical comments on many of these points, but some of the German comments were not accepted. A possible new revision of OECD 209 is not currently foreseeable.

This document provides answers to the questions most frequently asked in order to assure applicants and test institutes that the studies they perform will be accepted by the Federal Environment Agency in all relevant fields of enforcement (REACh, plant protection products, biocides, medicinal products).
The document describes how the Federal Environment Agency interprets the new guideline, how studies must be performed and in what cases deviations from some requirements of the Guideline are permissible.

Applicants should ascertain in advance whether studies so performed will also be accepted by the assessment authorities of other countries (particularly in the case of EU procedures).

The document will be updated as needed to incorporate other frequent questions or new findings.

A current version of the document is available at http://reach-info.de/pruefrichtlinien.htm.

Questions concerning the test guideline as well as comments on this document can be addressed by e-mail to Lars.Hohndorf@uba.de.

Is the determination of nitrification obligatory?

Testing for inhibition of nitrification was adopted from DIN EN ISO 8192:2007. The background is that in tests measuring total oxygen uptake rates, dose-response curves were sometimes biphasic or distorted or EC50 values too low. It was found that the activated sludge used in those cases nitrified significantly and that the test substance had a greater effect on the oxidation of ammonium than on heterotrophic carbon oxidation.
Paragraph 2 of the Guideline contains background information on why and when the effect on nitrification should be measured.

Measurement of nitrification is not obligatory. Rather, nitrification must be measured only if the activated sludge used is expected to nitrify significantly. This can be examined in a preliminary test using a nitrification inhibitor (also see paragraphs 2, 4, 6, 30, 38 of the Guideline).

Measurement of nitrification is also not required if, for example, no inhibition of total oxygen uptake was found in a preliminary test with the test substance.
Paragraph 4 'For most purposes, the method to assess the effect on organic carbon oxidation processes alone is adequate. (…)'
Paragraph 30 '…Normally, the measurement of total oxygen uptake inhibition should be adequate. …'

How many controls, concentrations and replicates must be tested and how did this requirement in the Guideline evolve?

It was only during the period for comment on the draft Test Guideline that demands to increase the number of concentrations and replicates were raised and discussed. The intention was to bring the new OECD Guideline 209 statistically into line with the existing ecotoxicity tests in order to ensure the reliability of determined NOEC and EC_x values.

The last sentence in paragraph 5 'Principle of the Test' states that the test is typically used to determine an EC_x of the test substance **and/or** the no-observed-effect concentration (NOEC). However, paragraph 41 of the Guideline only suggests a test set-up for the combined determination of NOEC **and** EC_x, consisting of 6 controls and 5 concentrations with 5 replicates each.
The statistical determination of NOEC and EC_x is known to require different test set-ups. During the discussion of the draft Test Guideline both test set-ups were included in the text which made the statistical guidance complex and difficult to understand. To avoid parallel testing of both test set-ups, the compromise of using just one to determine both NOEC and EC_x with sufficient reliability was found with the help of a statistician.

For combined determination, the number of controls, concentrations and replicates should be chosen as suggested by the Guideline. However, that test set-up can also be used for determining just the NOEC or the EC_{50} individually. In other words, that test set-up is independent of the endpoint needed.

On the other hand, it may be concluded from paragraph 41 that when determination of only a NOEC **or** EC_{50} is envisaged, deviation from this test set-up is possible as long as sufficient statistical reliability is assured by the requisite number of controls, test concentrations and replicates. This means, for the determination of the NOEC, normally fewer concentrations and more replicates and for the determination of the EC_{50}, more concentrations and fewer replicates will be used.
Corresponding studies will be accepted by the Federal Environment Agency.

For example, a test for EC_x with 5 concentrations (factor ≤ 3.2), 3 replicates each and, wherever possible, 6 controls is deemed to be an appropriate minimum test set-up, provided that it is ensured that the concentrations cover the envisaged range of effects.

Is it necessary in every case to determine a NOEC or may alternatives such as EC_{10} be used?

For the EU, the EC_x set-up must generally always be used. NOEC alternatives such as EC_{10} are acceptable. For non-European countries, before performing the study the applicant should contact the competent authority of the relevant country whether determination of a NOEC is required so as to choose the proper test set-up.

If no significant inhibition occurs up to the highest tested concentration, that concentration is deemed to be the NOEC (NOEC ≥ highest tested concentration).

Is it possible to perform a limit test and if so, what concentration must be used?

The Guideline allows a limit test with 6 controls and a test concentration with 3 replicates. It can be performed at a test concentration of 1000 mg/L, for example. However, if effects are measured in this test, a full new test with a series of concentrations must be performed. So, the Guideline gives latitude in choosing the test concentration in the limit test, but the chosen test concentration should be as high as possible (PEC/PNEC wastewater treatment plant). Assessment practice to date suggests that limit tests with a test concentration of 100 mg/L are recommendable.

Do the same requirements and criteria apply to all relevant fields of enforcement?

The same requirements and criteria apply to all relevant fields of enforcement (REACh, plant protection products, biocides, medicinal products).

Can results obtained for a test substance on nitrification inhibition in soil be transferred to wastewater treatment plants?

Results from soil micro-flora tests on nitrification inhibition by a test substance are not predictive for its effects on nitrification in a wastewater treatment plant.

When must chemical analysis be performed?

The Guideline gives clear information about when and how chemical analysis should be performed.
Chemical analysis is only necessary if the nominal test concentration in the test vessels is unknown. This is the case, for example, if prepared stock solutions of test substances have concentrations above the maximum water solubility and a fraction of the weighed-in test substance remains undissolved. In that case, it is not known how much of the test substance is actually transferred into the test vessel so that it is necessary to measure the concentration. To simplify matters, analysis should then always be performed before the addition of the inoculum.
Weighing the test material directly into the test vessel is generally recommendable for all test substances, particularly for problematic substances (poorly water soluble, adsorptive, hydrolyzing, volatile). This makes it possible to refer to the weighed nominal concentration, irrespective of whether the total amount of the test concentration is present in dissolved form, and is the most likely to reflect the situation in a wastewater treatment plant. Furthermore, it is known that effects may also occur at concentrations above maximum water solubility.

How must volatile test substances be dealt with?

The passages in the Guideline on that point call for an analysis to be performed in individual cases where test substances are unstable or volatile. This demand does not always seem to be appropriate.
Unstable and volatile compounds should also be weighed directly into the test vessel, and the nominal concentration should be referred to in their case as well. This avoids time-consuming and expensive analytics. The test substances should be transferred into the test vessels as the last step directly before the start of the test and as gently as possible. The test should be designed in such a way that volatilization is at low a rate as possible while ensuring at the same time a sufficient supply of oxygen. Since volatilization also occurs in sewers and wastewater treatment plants under real conditions, the test design should describe the “worst case” with regard to available concentrations in the test vessel so that the loss by volatilization in sewer pipes and wastewater treatment may be expected to be as high as or even higher than in the laboratory test.

A 3 h test period is also the minimum standard for volatile substances and also covers effects occurring in the first 30 min.

In addition, UBA has considered integrating trigger values for critical Henry constants, above which a modified test design would be required, as a recommendation into this Q&A document. First investigations by a test institute revealed, however, that the determined constants cannot representatively be applied to other laboratories, other test equipment or other test substances.

It is likely in practice that not all sewage sludges can meet the new validity criterion for oxygen uptake by the controls. Is this criterion based on scientific necessities?

The validity criterion for oxygen uptake by the controls was taken from DIN EN ISO 8192:2007 and has been a component of the Guideline since 1995. During the period for comment on the draft OECD Guideline in late 2009/early 2010, Germany had submitted critical comments on this point, but the German comments were not accepted.
Early test reports available to UBA reveal that some of the activated sludges used by test institutes meet the criterion easily whilst others fail to meet it. The experiences of test institutes show that activated sludge respiration rates differ significantly between wastewater treatment plants. In addition, measurements have revealed seasonal variations. However, there have been no observations to date which indicate that this has an influence on sludge sensitivity.

Since too-low rates of activated sludge respiration are just as unfavourable as excessive ones, it seems to be scientifically appropriate to have a validity criterion for this in order to identify atypical activated sludges from wastewater treatment plants that do not work properly or handle particularly high loads.
UBA believes that this validity criterion should be modified to reflect realistic values or a realistic range of values if the majority of the activated sludge samples used shows that the criterion is not met. UBA is examining in how far activated sludge in other countries and regions fails to meet the criterion in order to then suggest an appropriate revision of OECD Guideline 209.

Pending the outcome, the Federal Environment Agency will consider on a case-by-case basis whether it can accept studies in which the validity criterion has not been met, on condition that the sensitivity of the activated sludge to the reference substance is within the valid range and the study is valid and plausible in all other respects. Plausible reasons for the deviation from the validity criterion must be given by the test institute in the test report.

Applicants should ascertain in advance whether such studies will be accepted by the assessment authorities of other countries (particularly in the case of EU procedures).

The new Guideline provides that the coefficient of variation of oxygen uptake rate in controls should not be more than 30% at the end of the test. Isn't this value far too high?

This value does seem very high and studies with a coefficient of variation of 30% in controls might be difficult to evaluate. However, when the Guideline was drafted, values from ecotoxicity tests (e.g. OECD 201 35% section by section, OECD 211 ≤ 25%) probably served as a starting point.

It is recommended that when preparing test vessels, 3 controls should be set up at the start and 3 controls at the end of the measurement series and it should then be checked, using the log values (normal distribution) in a t-test for example, whether their mean values differ significantly. Subsequently, all controls can be pooled and the coefficient of variation be calculated.

The new validity criterion is not comparable with the previous one, which allowed an absolute deviation of not more than 15% between the control at the beginning and the control at the end of the measurement series.

Can paragraph 34 of the Guideline be understood to mean that 2 controls are sufficient when test vessels are prepared and analysed in parallel/simultaneously?

No, a total of at least 6 controls is also necessary, for statistical reasons, when test vessels are prepared and analysed in parallel/simultaneously. If test vessels are measured in batches simultaneously, at least two controls must be included in each batch, which means that a total of more than 6 controls may be necessary depending on the number of batches. Due to the large number of test beakers, sequential preparation and measurement will likely be the rule.

It should be noted in this context that the first sentence of paragraph 34 is wrongly worded. It should read: "Blank controls (F_B) have to be prepared at the beginning and end of the **test series** in tests in which the test beakers are set up sequentially at intervals."

Why must the statistical soundness of the results be ensured with so many controls, test concentrations and replicates if results may be expressed in "order of magnitude"?

The draft Guideline included the option of expressing results in "order of magnitude", and even critical comments by UBA could not fully remove this option. Since individual values and not value ranges are needed for risk assessment, the Guideline now allows results be expressed in order of magnitude only additionally. Otherwise the statistical effort would not be justifiable.

Must activated sludge always be exposed to the test substance for a period of 3 h?

The activated sludge exposure period must be at least 3 h. Measurement after 30 min is allowed only additionally. Some test reports available to UBA where measurements were carried out after 30 min and after 3 h show that determined effect concentrations may differ significantly. Also, 30 min is not representative of residence times in wastewater treatment plants.
Studies in which exposure periods are exclusively less than 3 h are no longer accepted by the Federal Environment Agency.

Depending on the expected effect of a test substance it may be appropriate in individual cases to choose longer exposure periods to more closely simulate real residence times in wastewater treatment plants and cover effects that might occur after a longer period.