Clarifications needed on full submission package

TST

1. **Section 1.1.1, page 5, Application 1-Grouping**: The text says that the "bioelution data are used to group substances of a given metal and predict if they will have similar or different oral bioavailability". On page 8 (section 1.1.2) the text says that the data generated by bioelution "is intended to compare the bioaccessibility of a given metal between different metal-containing materials […] in application 1-Grouping and to calculate the bioaccessible concentration of metal ingredients in alloys in application 2-Alloy classification". The sentence on page 5 relates bioelution data with prediction of bioavailability while on page 8 the outcome of the bioelution tests is clearly link to bioaccessibility. Please revise and make the text consistent.
   Response: the text was clarified.

2. **Section 1.1.5.1 c), Quality criteria applied to the test system, page 14:** The quality criteria for the test system and the negative control acceptance criteria refer both to the measurement of elements of interests in the simulated gastric fluid. The difference seems to be that in the case of the negative control vessel, the measurement is done at the end of the 2h incubation period and the value should be < 20 x SD of blanks of the lab, while in the case of the test system the measurement is done at time 0 and no elements of interest should be found or with an insignificant amount. Please confirm and clarify it in the text.
   Response: text was clarified. We removed the reference to "test system should be free (or with insignificant amount) of the elements that need to be investigated". This is already covered by using high purity reagents to prepare the 'test medium' and then analysing it after 2 hours (negative control), and by using new test vessels or pre-treating the glassware in nitric acid. If the negative control analysed after 2 hours satisfies the quality criterion (see below the current text in SOP), it can be assumed that the quality criterion of the test system has also been met.

> **6.7 Data Acceptance Criteria**
>
> A **test can be assumed valid (i.e. satisfies quality criteria of test system) if:**
>
> - The values of pH and temperature (T) of the test medium at the start of the test fall within the set boundary conditions: i.e. pH 1.5 ± 0.1 and temperature at 37 °C ± 1 °C.
> - The values of pH and T of the negative control at the end of the test fall within the set boundary conditions: i.e. pH 1.5 ± 0.1 and temperature at 37 °C ± 1 °C.

3. The three bullets of the quality criteria for the test system should be combined in one sentence:

   The test system should be free of the elements to be investigated (or with an insignificant amount) at an initial pH of 1.5 ± 0.1 and initial temperature of 37°C ± 1°C.
   What an insignificant amount means? This is too vague for quality criteria. Could an upper cut-off limit be established/provided?
   Response: as indicated above this text was deleted. Text was clarified to indicate that only pH and temperature are measured initially, time=0; while pH, temperature and metal release (test medium negative control) are measured at the end of the incubation period. The measurements of pH and temperature are part of the quality criteria of the test system (section 1.1.5.1 part c); the measurement of metals in negative control is part of the acceptance criteria applied to the results (section 1.1.5.1 part e).

Negative control: A test vessel with test medium only, with or without a clear resin in the case of embedded massive samples. A negative control run in parallel helps to detect background values of the elements of interest in the test fluid and helps rule out contamination. The measurements of the elements of interest in the negative control should generally be < 20 x SD of negative controls for the lab (based on 10 or more replicates). Higher values would indicate possible contaminations of the test system, glassware, etc.

4. **Section 1.1.5.1 d) and e), page 15:** The term "benchmark material" is used both in section d) linked to the assessment of transferability of the test method and in section e) as part of the test method to assess acceptance criteria. To avoid confusion, it may be better to use the term proficiency chemical for those materials used to assess transferability of/proficiency with the test method. Proficiency chemical is the normal term used in this situation in OECD Test Guidelines.
   Response: text was clarified by defining 'proficiency materials' as those that can be used to assess transferability. Proficiency materials are well characterized and defined materials (e.g. regarding elemental composition and particle size) that may or may not include the metal of interest but need to match the physical form of the test materials (e.g. samples of powder metals or metal compounds, embedded and non-embedded massive forms of alloys). These materials are tested initially by the lab when establishing the SOP and are run periodically afterwards (at least once a year). Proficiency materials are independent of the application of the test data. A physical repository and Shewart charts with test results will be made available. TST text and attachment 19 were edited. Furthermore, the term "benchmark materials" is not used anymore in the SOP or TST.

   Furthermore, the SOP now includes a positive control to be run in parallel with the test samples. Positive control: The positive control can easily be prepared, for example, by spiking certain amounts of the Certified Standard Metal solution into the test medium. Its purpose is to assure that the gastric test produces repeatable results. It provides evidence that the test system is responsive under the actual conditions of the assay. The positive control is run in parallel with the test materials of that metal or element. The positive control is independent of the application of the test data. The positive control does not need to match the physical form of the test materials. The positive control is predicted to give reproducible results over time. The positive control can also be used to assess the recovery of the metal ions through the test (mixing with test medium, incubation, filtering, sampling, etc). This solution shall be run at the time when the SOP for one or several metals of interest is established in a lab. Using this solution, high and low metal ion concentration would be added to the test medium at time = 0 and run through the protocol. Recovery of 100 ± 10% would be considered acceptable.

5. The text under d) says that measurements of elements of interest should have a recovery of ± 20% of target value. What it is meant by target value? Could these be provided? Later on, under section e) accurate metal measurements ± 20% of reported values from a benchmark material is included as acceptance criteria for the results. It is not clear if these are different criteria or the same one? It is also not clear why the option to run the benchmark simultaneously or periodically in the lab is provided. In our view, it should be run always in parallel. In case periodically is kept, please justify it and indicate how often a lab should test the benchmark material.
   Response: text was edited to indicate that a physical sample repository of the proficiency materials will be set up and a publicly available database (with Shewart charts) will be created to house the repeated measurements of the proficiency materials. Ideally, as tests are conducted and more measurements are added, the upper and lower CI for each proficiency material can be accurately determined in the Shewart charts. Initially, the database will have at least 10 measurements for each proficiency material collected by at least 2 different labs (see Annex 2 to SOP and Attachment 19 to TST). The text was edited to be consistent; e.g. 'accurate metal measurements ± 20% of reported

values in database. Regarding when a proficiency material should be run it was clarified that after the initial adoption of the SOP when the proficiency material will be run for the first time, the proficiency material could be re-run at least once a year.

6. *Attachment 19* includes a long list of "benchmark materials". It is not clear why so many are included and how the selection of the benchmark will be done for a particular experiment. Is this based on the element(s) of interest? This should be better described in the protocol.
Response: The test was clarified to indicate that the selection of the proficiency materials ***will be driven by the physical form of the test samples***. In discussions with the labs currently performing bioelution tests, it became clear that the proficiency materials do not have to contain the same elements present in the test material samples. The labs run many different samples and consider that the main challenges are associated with the physical form of the samples rather than the measured metals. It should also be noted that positive controls of each element of interest will be included in each run. Attachment 19 and Annex 2 to the SOP have been revised.

7. *SOP (Attachment 1a)* section 2.10 also mentions the benchmark material that can be used to justify the use of an analytical technique different from the ones recommended, by showing that the alternative technique is able to accurately measure the element of interest in benchmark materials. It is not clear against which values they will be compared. Are those for instance, historical release values for the element of interest in the benchmark material? If so, shall they be generated by the lab or are they reported somewhere for each benchmark material? Please clarify.

Response: text was edited to indicate that a publicly available database will be created to house the repeated measurements of the proficiency materials; these measurements will be obtained by at least two independent labs following the final approved SOP. Please see response to Q5. The text of TST section 1.1.5.1. (b) and the SOP have been revised as shown below.

6.4 Sample analysis

The concentrations of released elements of interest are measured with an appropriate (validated) measurement method. In order to guarantee the basic quality (use of a validated measuring method, involvement of appropriately trained staff, traceability of the measurements, data processing and archiving,…) the sample analyses are preferably performed in a laboratory that works according to a standardised quality system (e.g. GLP, GMP, ISO17025, ...).

Examples of universal and commonly used method for measuring dissolved element concentrations (and thus recommended) are:

- ICP-MS (inductively coupled plasma-mass spectrometry); well documented in ISO 17294-1:2004 Water quality - Application of inductively coupled plasma mass spectrometry (ICP-MS) - Part 1: General guidelines and ISO 17294-2:2016 Water quality - Application of inductively coupled plasma mass spectrometry (ICP-MS) - Part 2: Determination of selected elements including uranium isotopes.
- ICP-AES (inductively coupled plasma-atomic emission spectrometry); well documented in ISO 11885:2007 Water quality - Determination of selected elements by inductively coupled plasma optical emission spectrometry (ICP-OES)

However, other validated analysis methods can also be applied (such as AAS (graphite furnace or flame - atomic absorption spectrophotometry).

Unless otherwise agreed, only 1 measurement result is reported for each analytical sample and element (in µg/L or mg/L).

In addition to the measurement results of the test samples, a reporting limit is always provided for each element of interest (limit of quantification, LOQ or at minimum the limit of detection LOD, see glossary for definitions). Other analytical performance characteristics of the analysis (measuring range, repeatability, reproducibility, accuracy, measurement uncertainty) should be reported.

8. **Section 1.1.5.1 e), Experimental system, page 15:** The criteria listed here for pH and temperature are the same as the ones already indicated under c) for the test system. Since section e) refers to the results, should it not be more appropriate to include as criteria for pH the one indicated in footnote 16 (page 47), e.g. "end-of-test changes in pH should not be higher than 2-fold"?
Response: Thanks for this comment. The discussion on pH has been moved under section c) (*Quality criteria applied to the test system*). We had some additional discussions and have now edited the text as follows: "At the end of the test, the pH of the triplicate test item vessels may deviate significantly from 1.5, depending on the nature of the test material (e.g. tin) and should be recorded. In such a situation, the pH of the triplicate samples test material deviating significantly, it is recommended to compare the changes in pH between the high and low loadings. If the pH is above 4.5 and there is a difference between the low and high loadings, it shall be considered to use the results from the loading (expected to be the lower one) at which the pH is closer to 1.5. If the pH among the 3 test material replicates at the end of the test differ by more than 10%, contamination could be suspected and the test should be repeated."

9. It is not clear what exactly "visual observations should match/support the measurements" refers to and how this information will be used? Do you mean observation of precipitates?
Response: text was unclear indeed. Text was edited to indicate that any visual observations of e.g. precipitation, colouration, decolouration that differ between replicate test vessels or are consistent for all three replicates, should be recorded as they can help explain the bioelution results (section 1.1.5.1 e).

10. **Section 1.1.5.1 e), Performance of the analytical equipment, page 15:** From the analytical point of view, results must be obtained following a sample sequence that includes QC in between and carry over checks. The calibration of the equipment needs to be done using the same test system (i.e. simulated gastric fluid) as the one used with the samples. Is this assumed to be covered if an appropriate analytical technique (e.g. according to ISO 17025 or equivalent) is used as indicated in TST and protocols? Please clarify.
Response: The discussion on performance of analytical equipment was moved under section 1.1.5.1 b). Gastric fluid test system is fairly simple. After the filtration step, nitric acid is added prior to element analysis to preserve the metal ions in the solution; this is specified for example under *ISO 17294-2:2016: Water quality - Application of inductively coupled plasma mass spectrometry*. The Cl^- or NO_3^- ions are not expected to interfere with the analysis of metals. Thus, the analytical calibration (e.g. ISO17025) is considered sufficient to guarantee the accurate measurements of the released metals.

However, in order to test the recovery of metal ions of interest through the whole bioelution test (including filtering step), we propose to use the positive control (which can easily be prepared by spiking certain amounts of the Certified Standard Metal Solution used for the analytical into the test medium). The positive control solution shall be run at the time when the SOP for one or several metals of interest is established in a lab. Using this solution, high and low metal ion concentration would be added to the test medium at time = 0 and run through the protocol. Recovery of 100 ± 10% would be considered acceptable.

11. Please provide the data indicated under footnote 11 (i.e. measurements of Ni and Co samples) to support the assessment.
Response: The data is shown in table below. The tests were conducted at different facilities. This table is now included in the text of the TST (section 1.1.5.1 b).

Sample	Co (µg/L)		Ni (µg/L)	
	GAAF	ICP-MS	GAAF	ICP-MS
1	< LOD (0.3)	0.05	<LOD (1.0)	0.30
2	<LOD (0.3)	0.05	<LOD (1.0)	0.33
3	0.5-<LOQ (0.6)	n.m.	36620*	32496
4	0.5-<LOQ (0.6)	n.m.	36180*	31954
5	5.55	5.89	130.0	123.3
6	5.50	5.96	111.4	122.4

*AAS-flame; n.m. not measured

Sample 1 and 2: gastric fluid medium (negative control); samples 3 and 4: 0.2 g/L Ni powder (high Ni release); samples 5 and 6: SS316L powder (medium Ni release, low Co release)

12. **Section 1.1.5.1 e), blank values, page 16**: Are the method blank values the ones mentioned on page 14 as the blank control (= negative control) measurements (i.e. containing only the test solution)?
Response: yes, text was clarified by using exact same wording "negative control."

13. **Section 1.1.6, page 22, and Attachment 2**: Several references are cited, some unpublished. However, it would be useful to add another column indicating the gastric fluid used (i.e. the composition) and other relevant aspects of the method such as agitation, etc.
Response: two columns were added: one with composition and pH of gastric fluid and a second one with loading and duration of exposure.

14. **Section 1.2.2, 5th paragraph on ratio S_R:S_r, page 29,:** Section 1.2.2 is a mix of WLR and BLR. The section should describe WLR while the ratio S_R:S_r characterises BLR. More confusing is then the next to the last sentence stating that the fact that nearly all metals/substances had ratios < 6 demonstrates good within-lab agreement. The same sentence appears under 1.4.1, page 33, 4th paragraph, indicating in this case good repeatability and reproducibility.
The next paragraph also refers to both WLR and BLR.
Response: you are right. We have now edited the text to limit the discussion to within lab repeatability (WLR).

15. **Section 1.2.3, outlying values, page 29:** Section 1.1.8 shows values from Henderson et al, to discuss non-qualified tests. According to the information provided there, all 5 laboratories reported measurements > LOD (no values < LOD), while section 1.2.3 indicates that for Cr some labs reported values < LOD. Please clarify.
Response: you are right. The text in section 1.1.8 has been clarified to indicate that the analyses are based on the results for those samples that were part of the round robin (9 metal-sample combinations). For the 10th metal-sample combination (excluded from round robin data analysis) there were two labs that had 2/3 and 1/3 of the triplicate measurements < LOD. Text was also clarified in section 1.2.3.

16. **Section 1.2.4, last bullet referring to "synthetic perspiration treatment condition", page 30:** since the proposed test is designed to obtain results using synthetic gastric fluid, reference to other synthetic fluids that were, nevertheless, used in the round robin exercise may lead to confusion.
Response: agreed. Text was deleted.

17. **Section 1.2.4, pages 30-31:** the paragraphs quoting reports of Labs B and C refer to an annex and an

appendix not included in the submission package. This could lead to confusion if one wants to find them among the attachments provided with the TST. I may therefore be best to delete the references to these documents from the quotations included in the TST or otherwise, to include them in the submission package.
Response: agreed. Text was deleted.

18. **Section 1.4.2, page 34:** are the recommendations implemented in the SOP the ones included under section 2.1.6? If so, please refer to section 2.1.6. If not, please list the recommendations that were implemented.
Response: No. During the preparation of the manuscript (Henderson et al 2014) reporting on the results from the round robin study, a number of aspects of the various bioelution test protocols were identified that were thought to contribute to the increased variability observed with some fluids (e.g. bubbling of CO_2). Subsequent tests conducted by metal commodity associations using these fluids incorporated these lessons to improve repeatability and reproducibility of the tests. We have removed the text under Section 1.4.2, since this was not specific to the gastric fluid and could create confusion. The specific changes implemented in the current SOP, based on lessons learned from the round robin (gastric fluid) are listed in Table 9, section 1.6.2

19. **Section 1.5.2., page 39**: Table 8, PSD column: Why is the particle size distribution for SS316L shown at 80%, while for the mixture and Ni metal it is shown at 90%? In the ECTX report D90 for SS316L is reported as 6.2 µm. With regard to the Fe-Cr-Ni powder mixture, the ECTX report indicates that particle size distribution is < 44 µm with a note referring to constituent elements. From where the D90 = 2 µm of the powder mixture is coming from?
Response: The D90 for the SS316 powder should be 6.2 µm. Table 8 has been corrected. With regards to the Fe-Cr-Ni powder mixture, the ECTX report indicates that particle size distribution of the mixture of Cr, Fe and Ni powders is < 44 µm. The results reported in Table 8 pertain to the nickel releases. Since the pure Ni powder that is present in the mixture has a D90 = 2 µm, it seems reasonable to assume that the same particle size will be present for the Ni powder within the mixture.

20. **Section 1.6.1, page 41:** It is indicated that some bioelution tests have been conducted after 2012 with epoxy-embedded massive forms of alloys. Please refer to the study carried out by ECTX (X02a-146) and any other published work, if any.
Response: Text was edited to include references and additional results were added. The tables below were inserted under Section 1.1.8 (page 26). The first table includes results from triplicate measurements of two different alloy samples and two different metals (Ni and Cu). Results show good repeatability with one exception (CV 14%). The second table includes the results for Cu release from alloy CW602N when tested in the same lab (CIMM, Chile[1]) in 2010 and in 2015, and in a second lab (ECTX, Belgium) in 2018. When comparing results at the same loading (2 g/L) for the two labs, results show good reproducibility with values of 203 and 262 mg Cu/g massive alloy sample.

Variability in metal measurements from epoxy embedded massive samples
In all cases 3 independently prepared samples of the material were tested

[1] Personal communication from S. Baken of EU Copper Alliance.

A. Repeatability

	Samples and metals measured (µg /L) at a loading of 2 g/L				
	CW602N alloy 1h-Cu	CW602N alloy 2h-Cu	CW602N alloy 4h-Cu	SS316L 2h-Co	SS316L 2h-Ni
Replicate 1	177	385	966	0.126	2.94
Replicate 2	175	385	995	0.122	3.1
Replicate 3	224	447	1092	0.133	3.28
Mean	192	406	1018	0.127	3.11
CV%	14*	9	7	5	5
Reference	ECTX X02a-146			ECTX X02a-081	

*Does not pass the repeatability < 10% criterion

B. Reproducibility

	Mean 2 h copper values for alloy CW602N from 5 assays conducted at different loadings and at 2 different labs over an 8-year period				
Year of study	2010	2010	2015	2015	2018
Loading (g/L)	0.2	2	0.25	3.9	2
Metal release: µg copper/g alloy	217	262	462	229	203
CV%	5	3	19*	7	9
Reference	CIMM, 2010	CIMM, 2010	CIMM, 2015	CIMM, 2015	ECTX X02a-146

*Does not pass the repeatability < 10% criterion

21. **Section 1.6.2, page 41**: Is Ag the only metal outside the applicability domain? Several metals are discussed in the reply to questions Q1, Q4, and Q7 raised in the test pre-submission assessment report (Mo, Si, Sb, Ag, W, Hg) for which the release is pH dependent. Please clarify and align the text.
Response: Thanks for raising this point. It was confusing. The text was edited in this section and in reply to PARERE questions Q1 and Q4 to be consistent.
In addition to Ag, there are other elements (Mo, Sb, Si, W, and Hg) for which there is insufficient evidence to conclude that gastric fluid will yield the highest metal release values for the oral route. For these metals and metal compounds it is also recommended that in addition to the gastric fluid (current SOP) bioelution tests be conducted in saliva and/or intestinal fluid for oral route assessments under applications 1 or 2.

22. **Section 2.1.1, page 44 and Attachment 13:** Both the TST on page 44 and the Attachment 13 refer to an independent third party who supervised the overall study coordination; however, the identity of this third party is not disclosed. It is also mentioned that the QA of the reports generated by each laboratory was done by the study coordinator. Since Attachment 13 names three study coordinators (Dr. Oller and Dr. Henderson from NiPERA and Dr. Verougstraete from Eurometaux), it is not clear who from the coordination board was responsible for the QA exercise. Please clarify.
Response: Indeed. This was confusing. The text has been edited (TST and Attachment 13) to clarify the following steps:

- Prior to conduct of the round robin, an independent third party (T. Brouwers of ECTX) was asked to collect and distribute the samples to the initial 4 participating labs. Eventually the ECTX lab agreed to also participate in the round robin at no cost; thus 5 labs were included.
- The metal commodity sponsors were represented by R Henderson (NiPERA, who later moved to ToxStrategies and was replaced by A Oller), F Van Assche (International Zinc Association), K Delbeke (European Copper Institute), T Brock (Cobalt Institute) and V. Verougstraete (Eurometaux).
- R. Henderson (who by then had moved from NiPERA to ToxStrategies) was contracted as study coordinator to review the reports (conduct Q&A), coordinate the statistical analyses and prepare the manuscript for publication. V. Verougstraete and A. Oller provided advice.
- Two independent statistical analyses of the bioelution data generated by the round robin were conducted by P Wilrich (Institut fur Statistik und Okonometrie Freie Universitat Berlin) and ToxStrategies

23. **Section 2.1.3, page 45, and Attachment 4:** S_r and S_R values shown in Attachment 4 have been extracted from Henderson et al. 2014 (Table 4 "Results of repeatability and reproducibility analyses") but reported with a different numbers of decimals. Could you please explain why the values in the paper appear with 3 decimals but were rounded up to one decimal in Attachment 4? Could the reason be that these values are not precise enough to report 3 decimals? Nevertheless, the 0.0 values reported in Attachment 4 are difficult to interpret.
    Response: For consistency, we edited Attachment 4 to reflect the same number of decimals as in the Henderson et al. (2014) manuscript.

24. **Section 2.1.4, page 46:** In the round robin exercise all test materials were powders and were tested as received without further grinding or other manipulation to alter particle size. This information needs to be included since it is relevant with regard to the study conclusions in section 2.1.5. The possible impact of manipulation of the material to get an adequate sample size has not been assessed during the round robin. Similarly, the few data available with epoxy embedded massive forms of alloys (ECTX study X02a-146) does not allow making overall conclusions regarding variability/reproducibility of bioaccessible data and transferability of the methods to other labs. The conclusions made refer to powder samples that do not require further manipulation to get an adequate sample size. This needs to be clarified in the TST.
    Response: This is correct. We edited the text on sections 2.1.4 and 2.1.5 to include a specification that the samples included in the RR were tested as received and that modified samples or epoxy samples were not part of the round robin.

25. **Section 2.1.4, page 46:** Attachment 14 should read Attachment 3.
    Response: Text was corrected.

26. **Section 2.1.5, page 47**: It is claimed that the test method is easily transferable to other laboratories; however, module 3 (transferability) has not been carried out and therefore it is difficult to understand the evidence and/or criteria used to arrive to such conclusion. Please elaborate further and provide justification.
    Response: This may be a question of interpretation. We understood transferability as the ability of a naïve lab to take the gastric fluid SOP, apply it to a given test material, and obtain reproducible results compared to an experienced laboratory. We believe that the round robin study provides indirect information on this issue. Some of the participating labs had experience with a "gastric fluid bioelution test" and some did not (they only had experience with T/Dp testing). All the labs tested

the same samples of the test material using the same SOP and the results showed good reproducibility. We have edited the text to make this point clearer and to state that we have not carried out a stand-alone transferability study.

27. **Page numbering**: From page 29 onwards, the total number of pages is wrong. Please correct.
Response: This has been corrected.

Protocols

28. The protocol has been submitted in four different forms, three of which were in the form of a DB-ALM protocol or Method summary. Some discrepancies have been identified between the different versions. Rather than trying to harmonise the entire versions, at this stage we propose to focus on the protocol submitted in the form of a SOP (Attachment 1a) and the DB-ALM Protocol (Attachment 1b long).
Response: we have done so thank you

Please, harmonise both versions taking into account the comments below:

SOP in Attachment 1a

29. The SOP as described in Attachment 1a would benefit from a section on the purpose and applicability, describing to what the procedure applies and when it can be used. Text on the test material in section 2.1 from the first and last paragraph, describing that it applies to metals, inorganic metal compounds or complex metal-containing materials, of specific dimensions and not to nanomaterials, would fit in such a section. The purpose of testing is for instance described in the DB-ALM Method Summary (under Test Method definition).

Response: Thanks, this was done. The purpose is described in SOP section 1 and the applicability in SOP section 2 as well as in Annex 1 where examples of calculations are included.

30. Harmonise terminology: test material or test item, test medium or test system.
Response: We edited the SOP to use the term 'test material' to refer to what is tested, 'test medium' for HCl solution, and 'test system' for test medium + test vessels.

31. Explain abbreviations and terms in a specific section
Response: A glossary section has been added to the SOP.

Introduction

32. The paragraph in the introduction that refers to the use of the data from bioelution methods in regulatory decisions should be described elsewhere so as to keep the text in the SOP limited to the generation of data and critical aspects to consider when performing the experimental work.
Response: this part of the text has been removed.

Section 2.1 Test material

33. The first paragraph in this section is repeated in Annex 1 under Test materials. It is suggested to move the text in italics to Annex 1 and avoid repetitions. Furthermore, the SOP should make clear which minimum and maximum sizes can be tested. Maximum is 1 mm diameter, but what about the minimum (0.1 µm, which is the upper size limit for nanoparticles?)? Definition of nanomaterial and microparticles are included in paragraph 2 of Annex 3, which is provided as part of the reply to question 5 of the pre-submission assessment report. Considering the information extracted from the World Heritage Encyclopedia, should the minimum actually be set to 0.3 or 0.7 µm?

Response: this has been corrected in PARERE Annex 3. In the context of the SOP, for materials present in the market as powders (< 1 mm diameter) but not those in nano form (according to current EU definition of nanos, so (≥ 0.1µm). We recommend to test in addition the subfraction of particles < 100 µm diameter. These particles are relevant to the fraction of inhalation exposure that can be swallowed. We recognize that there may be cases where samples are placed on the market at a particle size range that may be >100 µm and <1 mm, with very little mass in the < 100 µm range. In this case the material should be tested as it is.

The filter has a pore diameter of 0.2 µm; thus particles (or complexes) smaller than that would be counted as free ion. Based on the review of oral uptake included as Annex 3 to PARERE, it is clear that small particles (<0.2 µm) can actually be absorbed whole through the intestinal lining (enterocytes can take up particles of 200 nm via endocytosis). Thus, we realized that we were putting too much effort into making sure that particles < 0.2 µm are not counted as soluble ions when in practice they should be counted if they can be absorbed as such. This is now noted in the SOP as shown below.

> Note: The effectiveness of the filter to separate dissolved from undissolved metal for a particular test material can be confirmed by visual inspection and/or light scattering or equivalent method. Any particles <0.2 µm can in theory pass the filter and will be counted as soluble ions. This could result in overestimating the bioaccessibility of the metal from the sample (conservative approach).

While still excluding pure nanos from the applicability of the current SOP, we can accept that any particles that get through the filter (be it a complex or a particle) be counted as "soluble" since we are using relative oral bioaccesibility as a conservative estimate of relative oral bioavailability. Regarding nanos, we need to further monitor what is discussed in the nanos' debates regarding its definition and ensure consistency.

Section 2.9, filtration equipment

34. Why is filtration (0.2 µm filter) used instead of centrifugation of samples? In fact, centrifugation is mentioned as an option when the filter is not enough to separate dissolve from undissolved. The text in Note 1 is rather general without providing clear criteria that allow checking if the filtration was successful. Some of the measures suggested are rather subjective (e.g. visual inspection). Please clarify.

Response: Filtration is the method used in the ASTM D5517. The Certified Standard Solutions for the element(s) of interest that is used during the analytical calibration and to prepare the positive controls, can be used initially (when a lab is starting to apply the SOP) to check the recovery of the metal ion of interest throughout the whole test system including the filter. This would be done by each lab for one or more metals once and would only need to be repeated if a change in equipment occurs (e.g., switch to a different brand of filter or test vessel). Recovery of 100 ± 10 % would be acceptable. Please see response to Q 10.

Section 2.10, Analytical instruments

35. ICP-OES is mention in TST. Is there a particular reason why in the protocol ICP-AES is mentioned?
Response: ICP-AES and ICP-OES are exactly the same technique by 2 different names. We have now consistently referred to ICP-AES in the TST, the DB-ALM and in the SOP (page 17), even if the reference ISO for measuring elements in water refers to ICP-OES.

- ICP-AES (inductively coupled plasma-atomic emission spectrometry); well documented in ISO 11885:2007 Water quality - Determination of selected elements by inductively coupled plasma optical emission spectrometry (ICP-OES)

Section 3.1

36. In addition to the standard ASTM D5517 'for determining extractability of metals from art materials', which is provided in the TST as Attachment 17 (references), the SOP also mentions the standards ASTM F963 for 'Consumer safety specification for toy safety' and EN 71-3: 1994 for 'Safety of toys'. As these standards are not available to EURL ECVAM, it cannot be assessed if the preparation of the simulated gastric fluid proposed in the SOP is done exactly as in the above-mentioned standard(s) or if differences occur. Please provide the standards ASTM F963 and EN 71-3: 1994 or consider removing their mentioning from the SOP.
Response: these standards are included in this submission.

Section 3.2

37. Some instructions in this section, describing the test conditions, are repeated in more detail in sections 3.3 and 3.4. Please consider to re-write section 3.2 to make it an overview of the method, without giving instructions already that are provided later.
Response: This section (now section 6) has been rewritten to include a "Sample loadings (6.1)-", "Test setup" (6.2)- with specific details depending on the type of material, and an "Observations, sampling and measurements at the end of the test (6.3" section). This avoided some repetitions.

38. With regard to the incubation time, even if not critical, it is better to specify a range within which the operator has the time to process the samples.
Response: This has now been clarified to indicate that the operator should not take longer than 10 min per test material to process the samples at the end of the incubation period.

Section 3.3.2

39. In the example of epoxy mounted massive test samples, a more detailed description of how to translate from mass loading to surface loading is needed. In this regard, it is suggested to move the example included as note 3 under section 3.2 to section 3.3.2. The ECTX report (study n° X02a-146), under section 7.2 preparation of test item, provides a clear description of the calculation of the loading as well as the whole sample preparation (cutting, embedding and surface preparation). Some of these steps are missing in the current SOP.
Response: additional paragraphs explaining the calculation of equivalent surface loading and

providing examples have been added to section 6.2. The steps (cutting, embedding and surface preparation) for the epoxy embedded sample taken from ECTX reports have been included (section 3.4)

Section 3.4

40. It is recommended that the note in section 3.4 on the Sampling that states 'Selecting a 'midstream' sample helps avoiding deviations due to possible absorption to the filter', becomes an instruction, for better standardisation, i.e. always pre-wet the filter with 1 ml of the sample.
Response: The text was clarified to indicate that the filter should be prewetted with 1 mL of gastric fluid (test medium) before filtration of test material to avoid metal adsorption to the filter.

Data analysis

41. According to section 1.1.3 of the TST, the test method should always be used in a comparative way and it is intended to determine relative bioaccessibility: for application 1 - to compare the bioaccessibility of a given metal between different metal containing materials (e.g. comparing target and reference materials) and for application 2 - to calculate bioaccessible concentration of metal ingredients in alloys. The protocol lacks a section that clearly explains how the raw data are processed for each application (as shown with the examples included in the TST under section 1.1.5.2).
Response: Annex 1 to the SOP describes how the bioaccessibility data can be used in Applications 1 and 2 and provides numerical examples.

Section 3.7

42. The acceptance criteria that confirm validity of the experiment (as listed in Section 1.1.5.1 of the TST) are insufficiently described in the SOP.
Response: this section (now section 6.7) was revised.

43. It is not clear whether control measures are in place to confirm that agitation, settling, sampling and filtration were successful so that any dissolved metal could have passed the filter when it was present in the sample. This is not described in the current version of the SOP. For example, an internal standard of a frequently tested and soluble metal can be added to the vessel, together with the test material, and a specific amount of this metal should be recovered in the sample.
Response: See answer to Q34 and Q10 about the use of a positive control prepared using the Certified Standard Metal solution to check the performance of the SOP.

44. The inclusion of benchmark or proficiency materials to assess the transferability of the protocol to the laboratory is not enough to confirm the validity of individual experiments.
Response: this has been revised. Section 6.6 describes the use of a positive control (containing the element of interest) that will be run in parallel with the test materials. This is consistent with section 1.1.5.1 d) in TST.

Annex 1

45. All instructions that must be performed by the operator (storage, handling, material description and treatment) are usually described in the core text of an SOP. For example, the preparation and description of material (now described in Annex 1) could be placed in the core text of the SOP, as it is an important step of the procedure. Nevertheless, if the Annex is kept, clear reference to this annex should be provided in the SOP.
Response: this has been addressed in the revised SOP version by moving the information into the main part of the SOP and eliminating this annex.

DB-ALM protocol_stand alone document (Attachment 1b long)

Page 6

46. Participation in validation study: Remove the sentence regarding the OECD Test Guidelines programme since that refers to regulatory acceptance of the method and not to its validation.
Response: this has been corrected

Page 9

47. Within the analytical techniques listed, ICP-AES appears while in the TST ICP-OES is mentioned. Is there a particular reason for that?
Response: this was corrected. ICP-AES and ICP-OES are exactly the same technique by 2 different names. We have now consistently referred to ICP-AES in the TST, DB-ALM and in the SOP, even if the reference ISO for measuring elements in water refers to ICP-OES.

48. Consider coping here the text from SOP section 2.10 (Attachment 1a) that refers to the need to justify if different analytical techniques from the ones listed are used.
Response: the corresponding section from the revised SOP has been included

49. Consumables: keep here just the list of consumables and move the text that explains how to prepare/clean the test vessels and how to check the effectiveness of filtration to other sections.
Response: This has been done.

Page 10

50. Last bullet point of MEDIA, REAGENTS, SERA, OTHERS refers to certified reference materials for each element of interest. Are those the benchmark materials mentioned in the TST and in Attachment 1a? Please harmonise.
Response: Test was clarified for consistency. A Certified Standard Metal solution is used and contains the metal(s) of interest (for analytical calibration and to prepare the positive controls).

51. Test compounds: the relevant information from Annex 1 of the SOP (Attachment 1a) with regard to sample handling and preparation of the test materials should be extracted and included here.
Response: This has been done

52. Add here the relevant text from the consumables regarding the test vessels.
Response: This has been done

53. Experimental system - Simulated gastric fluid: The table included in the SOP (Attachment 1a) includes NaOH solution that is not mentioned here. Please harmonise.
Response: Text was clarified for consistency (inclusion of SOP table).

Page 11

54. Sampling time under ROUTINE PROCEDURES: See comment above under Section 3.2 (Attachment 1a).
Response: text was edited.

55. The last bullet point under ROUTINE PROCEDURES mentions 2 loadings but this has not been explained before. Harmonise with the text in SOP - section 3.2 (Attachment 1a).
Response: text was edited for consistency with SOP.

Page 13

56. Acceptance criteria are missing for the experimental system and the performance of the analytical equipment.
Response: text was added.

57. Data analysis: see comment above on data analysis for the SOP (Attachment 1a).
Response: text was added.

Page 14

58. Remove reference given in brackets to the attachments of the TST.
Response: text was edited.

Replies to the questions raised in the pre-submission assessment report

Question 4

59. Figure 9 reporting values for Sb, refers to "unpublished" work (cited as Hedberg et al, 2010b), which is not provided with the full submission. Would it be possible to share this work?
Response: yes, we have now included the report (Hedberg Y, Jiang T, Odnevall-Wallinder I. 2010b. Bioaccessibility of antimony released from four different antimony compounds in synthetic biological media. Commissioned by International Antimony Association (i2a)) in the submission package (181109Attachment17rev).

60. Please check figure numbers on pages 11 (paragraph on antimony referring to figure 7, left panel) and 12 (figure 8 next to MoO_2).

Response: Figure numbers were corrected. Thanks

61. On page 12, it is believed that the paragraph that comments the results on iron and molybdenum releases shown in figure 10 refers only to the ones obtained with the simulated gastric fluid. If this is the case, please revise the text and make it clearer.
Response: Text was revised.

Question 7

62. On page 17, it is claimed that precipitation should not be an issue with regard to the proposed protocol. However, in section 1.6.2 of the TST, it is claimed that AgCl precipitates can be found in gastric fluid. Please clarify and harmonise.
Response: There are two separate issues discussed in these docs that we may not have clearly distinguished: a) precipitation or complexation that can be observed in gastric fluid in vitro and may also be found in vivo (precipitation in HCL can be an issue for Ag and Hg) and b) metal releases in gastric fluids that may be lower than in other oral fluids at neutral pH for metals such as Ag, Mo, Sb, Si, W, and Hg. A precipitation or complexation does not necessarily invalidate the gastric fluid test if it reflects the in vivo situation. Lower releases in gastric fluids indicate that bioelution tests in other oral fluids (saliva or intestinal) should also be carried out for applications 1 and 2. PARERE question 7 deals with issue (a) precipitation/complexation, while PARERE question 1 deals with issue (b) lower metal release in gastric fluid than other oral fluids.

Question 9

63. Table 7, page 21: The value reported for Ni metal release (163 mgNi/g sample) is not found in the ECTX, 2017 report mentioned below the table. Since ECTX final reports have been submitted recently, it is suggested to update the table with the values found in the report from 2018.
Response: Table was edited to match the results to the appropriate ECTX reports since the nickel metal sample has been measured a few times. The numbers in the table were updated with final numbers in ECTX reports.

64. Table 8, page 22: check the values reported in column % Ni released. Should the values for Ni metal + SiO_2 and for SS316L be 2.7% and 0.007%, respectively, rather than 25% and 0.06%?
Response: Yes. Table was edited and final numbers from ECTX reports were added. 2.7% is the percent of Ni released from the Ni metal + SiO_2 sample while 0.007% is the percent of nickel released from the SS316 sample.

65. ECTX report (X02a-144), version from April 2018: we believe that on page 19 (last bullet), 130 ± 1 **mg**/L should read 130 ± 1 **µg**/L. Should the 0.061% of Ni release on page 20 (first bullet) be 0.007%?
Response: Thanks. The units should indeed be µg/L. The ECTX report has been amended. The 0.061% of Ni release is correct. It represents the fraction of Ni in the sample that is released and not just the fraction of the total mass of the sample.

Question 15

66. Figure 18, page 31: the red circles are placed wrongly. The gastric fluid on the X axis is the last one from the gastric phase (more or less in the middle of the figure).

Response: Thanks. This was corrected.

Question 17

67. Could you please justify your claim that there is no need to repeat the benchmark for each experiment? See also our previous comment on Section 1.1.5.1 d), e) page 15 of the TST.
Response: The proficiency materials should be run initially when the SOP is set up in the lab to assure that the lab is able to process this type of samples accurately in the gastric bioelution test. For this reason, the proficiency materials do not necessarily have to contain the elements of interest for a particular run of test materials. Thus, the proficiency materials do not need to be run every time test materials are run. They could be run periodically (e.g. at least once a year). In addition, a positive control (including the element of interest) is now included with each run.