Inter-laboratory validation of bioaccessibility testing for metals

Citation	Henderson, R. G., Verougstraete, V., Anderson, K., Arbildua, J. J., Brock, T. O., Brouwers, T., ... & Oller, A. R. (2014). Inter-laboratory validation of bioaccessibility testing for metals. Regulatory Toxicology and Pharmacology, 70(1), 170-181. doi:10.1016/j.yrtph.2014.06.021
DOI	10.1016/j.yrtph.2014.06.021
Publisher	Elsevier
Version	Accepted Manuscript
Terms of Use	http://cdss.library.oregonstate.edu/sa-termsofuse

Interlaboratory Validation of Bioaccessibility Testing for Metals

Henderson, Rayetta[*,a], Verougstraete, Violaine[b]; Anderson, Kim[c]; Arbildua, José J.[d]; Brock, Thomas O.[e]; Brouwers, Tony[f]; Cappellini, Danielle[g]; Delbeke, Katrien[h]; Herting, Gunilla[i]; Hixon, Greg[a]; Odnevall Wallinder, Inger[i]; Rodriguez, Patricio H.[d]; Van Assche, Frank[j]; Wilrich, Peter[k]; Oller, Adriana R.[l]

*Corresponding author

[a]ToxStrategies, Inc.
9650 Strickland Rd., Suite 103-195
Raleigh, NC 27615, USA
Phone: 919-797-9938
Email: rhenderson@toxstrategies.com and ghixon@toxstrategies.com

[b]Eurometaux
Avenue de Broqueville 12
1150 Brussels, Belgium
Email: verougstraete@eurometaux.be

[c]Oregon State University
Corvallis, OR 97331, USA
Email: kim.anderson@oregonstate.edu

[d]CECM, Adolfo Ibañez University
Adolfo Ibañez University
Diagonal Las Torres 2640
Peñalolen, Santiago, Chile
Email: Patricio.rodriguez@uai.cl and jose.arbildua@uai.cl

[e]Duke University
2200 West Main Street, Suite 400
Durham, NC 27705, USA
Email: thomas.brock@duke.edu

[f]ECTX bvba
Havenstraat 46/0.01, B-3500
Hasselt, Belgium
Email: tony.brouwers@ectx.be

[g]Kirby Memorial Health Center
71 North Franklin Street
Wilkes-Barre, PA 18701, USA
Email: dcappellini@epix.net

[h]European Copper Institute
168 Avenue de Tervueren
1150 Brussels, Belgium
Email: katrien.delbeke@copperalliance.eu

[i]KTH Royal Institute of Technology
Drottning Kristinas väg 51, SE-10044 Stockholm, Sweden
Email: ingero@kth.se and herting@kth.se

[j]International Zinc Association
International Zinc Association
Avenue de Tervueren 168/Box 4
B-1150, Belgium
E-mail: fvanassche@zinc.org

[k]Freie Universität Berlin
Promenadenstr. 16 A
D-12207 Berlin, Germany
E-mail: wilrich@wiwiss.fu-berlin.de

[l]Nickel Producers Environmental Research Association, Inc.
2525 Meridian Parkway, Suite 240
Durham, NC 27713, USA
Email: aoller@nipera.org

Abstract (199 words)

Bioelution assays are fast, simple alternatives to *in vivo* testing. In this study, the intra- and inter-laboratory variability in bioaccessibility data generated by bioelution tests were evaluated in synthetic fluids relevant to oral, inhalation, and dermal exposure. Using one defined protocol, five laboratories measured metal release from cobalt oxide, cobalt powder, copper concentrate, Inconel alloy, leaded brass alloy, and nickel sulfate hexahydrate. Standard deviations of repeatability (s_r) and reproducibility (s_R) were used to evaluate the intra- and inter-laboratory variability, respectively. Examination of the s_R:s_r ratios demonstrated that, while gastric and lysosomal fluids had reasonably good reproducibility, other fluids did not show as good concordance between laboratories. Relative standard deviation (RSD) analysis showed more favorable reproducibility outcomes for some data sets; overall results varied more between- than within-laboratories. RSD analysis of s_r showed good within-laboratory variability for all conditions except some metals in interstitial fluid. In general, these findings indicate that absolute bioaccessibility results in some biological fluids may vary between different laboratories. However, for most applications, measures of relative bioaccessibility are needed, diminishing the requirement for high inter-laboratory reproducibility in absolute metal releases. The inter-laboratory exercise suggests that the degrees of freedom within the protocol need to be addressed.

Keywords: metals; alloys; UVCB; classification; bioavailability; bioaccessibility; read-across; inter-laboratory validation

Abbreviations:

CEN	European Committee for Standardization
CLP	Classification, Labelling And Packaging of Substances and Mixtures Regulation
RBA	Relative Bioavailability
ECHA	European Chemicals Agency
RBALP	Relative Bioaccessibility Leaching Procedure
REACH	Registration, Evaluation, and Authorization of Chemicals
RSD	Relative Standard Deviation
s_r	repeatability standard deviation
s_R	reproducibility standard deviation
UBM	Unified BARGE Method

1. Introduction (7100 Text words)

As the demand for understanding the potential hazard and risk of chemicals to human health continues to grow, the data required for elucidating these concerns continues to expand as well. Meeting the new and evolving demands of regulatory programs such as the Registration, Evaluation, and Authorization of Chemicals (REACH) Regulation in Europe (EU) (Regulation (EC) No 1907/2006, 2006) necessitates the generation of new and scientifically robust data on chemical substances, including metals. The *in vivo* testing that would be required to fill these needs is often cost-prohibitive and time-consuming, and also raises concerns with regards to animal welfare due to the extent of testing potentially required. As such, alternative approaches such as read-across (extrapolation of known data from one substance to another substance) based on structure activity relationships or bioavailability are often encouraged to perform hazard and risk assessment while reducing animal testing (ECHA, 2008; 2013). For most routes of exposure and health endpoints, it is indeed the bioavailability of the metal at the target site in an organism that is the most important factor determining its potential toxicity. Bioaccessibility, referring in this context to the amount of metals released from a given material in fluids designed to mimic those of the human body and may become available for uptake (e.g., synthetic gastric fluid to simulate oral exposure) (Ruby et al., 1999; Henderson et al., 2012), provides a conservative estimate of bioavailability. Bioaccessibility is measured in *in vitro* bioelution assays, whose application to hazard and risk assessment has been increasingly used as an alternative to *in vivo* testing in recent years. Bioaccessibility is a conservative concept because not all metals available

will be absorbed or induce damage (effects will depend on dose and metal speciation). Such data are particularly informative, as the presence of a metal does not always impart its biological properties on a given material, for example when the release of the metals and their absorption may be limited due to surface and material properties (e.g., for alloys).

The comparison of bioaccessibility data for two or more forms of the same metal (e.g., a pure metal and an alloy with the same metal constituent) enables an estimate of their relative *in vivo* bioavailability. This type of information can be used in a variety of ways for metals assessment, including: as a tool in determining hazard classification (e.g., using relative bioavailability to determine classification or justifying a derogation because of a lack of bioavailability; ECHA, 2013), to aid in establishing categories of metal substances (grouping; ECHA, 2008), as part of the weight of evidence approach applied in performing read-across (e.g., Henderson et al., 2012); and for risk assessments for exposure to metals required by some consumer product safety regulations (Brock and Stopford, 2003). In addition, relative bioaccessibility can be used to estimate the effective concentration (defined as the fraction of released metals in biological fluids compared with its matrix concentration) of a metal in a complex material where matrix effects may occur (e.g., alloys) and enable read-across between these materials (Stockmann-Juvala et al., 2013; Hedberg et al, 2013).

The bioaccessibility concept is already incorporated in some standard bioelution test methods and regulatory frameworks, such as the European standard for release of nickel

in artificial sweat (BS EN 1811, 2011), ASTM D5517 (2007) for metals in art materials, and EN 71-3 (2013) that specifies safety requirements for metals in toys. Bioaccessibility has been listed as a possible approach for complying with information requirements of REACH as part of the chapter on grouping of chemicals (ECHA, 2008).

Method development for – and utilization of – bioelution testing by independent and government research groups have increased in recent years. The bioaccessibility approach to estimate metal bioavailability has been applied in recent years to human exposures to metals and minerals in soils, consumer products, and to the evaluation of metal substances (Hillwalker and Anderson, 2014; Henderson et al., 2012; Stopford et al., 2003; Herting et al., 2008; Hedberg et al., 2010; Mazinanian et al., 2013; Oller et al., 2009; Hamel et al., 1998; Vasiluk et al., 2011; Drexler and Brattin, 2007; Wragg et al., 2011; Ellickson et al., 2001; Turner, 2011; Gray et al., 2010; and Twining et al., 2005; Hedberg et al., 2013; Hedberg and Odnevall Wallinder, 2013; Jiang et al., 2012; Hedberg et al., 2012). In addition, some groups have developed research programs to perform inter-laboratory validation of bioelution methods for specific systems and metals. For example, Drexler and Brattin (2007) reported the outcome of a validation exercise for a method to estimate *in vivo* bioavailability of lead (Pb) from soils. Additionally, a separate group also performed a round-robin study for a different physiologically-based method for estimating the bioaccessibility of Pb, as well as cadmium (Cd) and As, from soils (Wragg et al., 2011). Cordeiro and co-workers (2012) reported the results of an inter-laboratory comparison of 8 metals in comminuted flakes from alkyd resin paints simulating a toy coating using EN 71-3 (1994).

Although some groups have sought to standardize specific methods (Drexler and Brattin, 2007; Wragg et al., 2011; Ashley et al., 2012; Cordeiro et al., 2012), generally standardized fluid compositions and testing protocols for the basic bioelution method are lacking. In addition, there are no reference standards to ensure the accuracy of these bioaccessibility results and existing studies have demonstrated that sample characteristics and methodological differences (e.g., temperature, pH, sample loading) can affect the amount of metals released (Stopford et al., 2003; Midander et al., 2006; Hedberg et al., 2013).

The aim of the current study, therefore, was to perform a cross-laboratory testing of different metal-containing materials in select simulated biological fluids that are relevant to characterizing key routes of human exposure, using a defined protocol. To do so, five laboratories measured the release of metals from six different metals and metal-containing materials in synthetic gastric, lysosomal/interstitial, and perspiration fluids (representing oral, inhalation, and dermal routes of exposure, respectively). The results of these bioelution analyses were evaluated by characterizing within-laboratory repeatability and between-laboratory reproducibility measures.

2. Materials and Methods

2.1 General study design

The five laboratories participating in the inter-laboratory validation study were Center of Ecotoxicology and Chemistry of Metals, Universidad Adolfo Ibañez (Santiago, Chile), ECTX-Consult (Hasselt, Belgium) with analytical work conducted at Labtium Oy (Finland), Kirby Memorial Health Center (Wilkes-Barre, PA, USA), Oregon State University (Corvallis, Oregon, USA) and KTH Royal Institute of Technology (Stockholm, Sweden). Each laboratory was assigned an identification code of A-E in no specific order and is referred to by its respective coding throughout this manuscript. All labs performed bioaccessibility testing in the following four simulated biological fluids: gastric, lysosomal, interstitial, and perspiration. Labs were asked to follow a Standard Operating Procedure (SOP; dated November 2010) provided and discussed prior to study initiation. In brief, test materials were added to simulated fluids and extracted for a set period of time under standard conditions (e.g., pH, temperature). Following a filtration step, extracts were analyzed and the amount of metals released into solution was reported. Laboratories measured the release of seven different metals (Cr, Co, Cu, Fe, Ni, Pb and Zn) depending on the composition of the test materials.

2.2 Test Materials

The six materials tested are listed in Table 1 with their respective chemical formula, CAS number, metal content, mean particle size, surface area, and supplier. The materials were Co oxide, Co powder, Cu concentrate, Inconel alloy, leaded brass alloy, and Ni sulfate hexahydrate. All test materials were powders with a median particle size <60 μm in

diameter representing a size range relevant for oral and dermal exposures, and compliant with requirements of ASTM D5517 (2007) and BS EN 1811 (2011). However, although the SOP required particles sized <10 μm for testing in interstitial and lysosomal fluids, which is considered to be representative of the respirable fraction, only three samples met this criterion. As Ni sulfate hexahydrate is hygroscopic, the salt agglomerated to a mean particle size of 12.4 μm. However, its particle size is not relevant as it is readily soluble in aqueous solutions. The copper concentrate was ground during the concentration process and the smallest attainable particles were sent to the labs for testing (mean diameter of 59.2 μm). As lead in the leaded brass alloy sample has lubricating properties, additional milling would have likely smeared the particles together. Therefore, a sieve was used to separate the smallest fraction for testing with a mean particle size of 56.2 μm. Laboratories were supplied with 100g of each test material from the same original batch and samples were tested as received without further grinding or other manipulation to alter particle size.

2.3 Laboratory Equipment

In general, laboratories used similar equipment and any major deviations are listed in the Supplemental Online Material. All chemicals used to prepare the test fluids were of analytical grade reagent quality or better unless otherwise stated. Test vessels were inert, chemical resistant, covered Erlenmeyer flasks of 250 mL. All glassware was cleaned by acid soaking for 24h (10% HNO_3) then rinsed four times in ultrapure water (18.2 MΩcm) and dried (by air or oven). A thermostated linear shaker (37 ±1°C, 150

rpm; stroke length=1 inch) or a thermostated orbital shaker (37 ±1°C, 171 rpm stroke length=1 inch) was used for agitation. Controlled thermometers with a readability of 0.1°C and calibrated pH meters with a readability of 0.01 units were utilized. A calibrated micro balance with a readability of 0.01 mg or 0.001 mg was used. For filtration, 0.2 µm membrane filters. (e.g., Whatman UNIFLO syringe filters, Pall Acrodisc syringe filters or equivalent filter system), latex- and oil-free syringes, and polypropylene tubes were used.

2.4 Bioaccessibility Assays

All fluids and experimental set ups were prepared by each individual laboratory. The compositions and general testing conditions of each of the simulated fluids, including pH, temperature, loading, and extraction duration, are described in Table 2. The use of synthetic gastric fluid (pH 1.5) to represent oral exposure has been used extensively, starting with the Comité Européen de Normalisation standard, Safety of Toys (BS EN 71-3, 2013), which has been adopted in the United States as ASTM D5517 (2007; Standard Method for Determining the Solubility of Metals in Art Materials). Interstitial and lysosomal fluids are used as surrogates for inhalation. Interstitial fluid (pH 7.4), comprised primarily of Gamble solution, represents fluid deep within the lung and has been used for many years to evaluate a range of materials. In this study, 5% CO_2 in air was used to keep the interstitial fluid test solutions at pH 7.4±0.2. The approach used by each laboratory to maintain this pH varied and is described in the Supplemental Online Material. Simulated lysosomal fluid, which mimics intracellular conditions with a pH of

4.5 similar to that found in lysosomes of alveolar macrophages, was also used (de Meringo et al., 1994; Stopford et al., 2003). Finally, synthetic perspiration (pH 6.5) was used to represent release from test materials on the skin and was prepared according to BS EN 1811 (2011).

Ultrapure water was added to the fluid compositions listed in Table 2 up to a final volume of 1 L. Temperature and pH were measured at the start of each test and fluids were adjusted with HCl or NaOH as necessary to achieve the desired pH. Temperature and pH were also measured in the remaining blank control for each test solution after sampling. All bioaccessibility tests were conducted at 37°C except for tests in synthetic perspiration where a temperature of 30°C was used (BS EN 1811, 2011). Sample loadings were 0.2 and 2.0 g/L for gastric and all other fluids, respectively (Midander et al., 2006; Henderson et al., 2012; Stopford et al., 2003; Turner, 2011).

Extractions in gastric fluid were conducted for 2 h based on an average half time for gastric emptying of 17.7 min and complete emptying of 91 min in human volunteers (Tomlin et al., 1993; Wang et al., 2001). In addition, this duration has been shown to be correlated with acute oral toxicity of nickel compounds in a recent study by Henderson et al. (2012). All other extractions were carried out for 24 h or 168 h to be representative of longer-term exposures. All extractions were prepared and analyzed in triplicate.

Filtered extracts from blank controls and test vessels were analyzed for metal concentrations using ICP-OES, ICP-MS, or AAS (flame or graphite furnace, depending

on concentration) as noted in the Supplemental Online Material. Bioaccessibility measurements underwent a Quality Assurance (QA) check and were reported as released μg metal /g sample.

2.5 Quality Assurance

Each laboratory generated a comprehensive report, which underwent a QA exercise. A detailed review and comparison between the SOP and the 5 laboratory reports was performed. As part of this review, individual exchanges were held with the labs to address information gaps and confirm data when necessary. Some differences in methodology between labs were noted. As a result of this exercise, some datasets were excluded from statistical analysis.

2.6 Statistical approach

Amounts of released metals that were not reported by the laboratories or were below the respective limit of detection were excluded from any analysis. In addition, any fluid/time point/ lab dataset with 2 or more labs reporting results <LOD were excluded from the inter-laboratory validation.

The statistical analysis of the measurement results was based on ISO 5725-2 (1994). According to this method, measurement results obtained in an inter-laboratory study are inspected for consistency by plotting Mandel's h and k statistics and for outliers by

application of the Grubbs tests and the Cochran test. A laboratory mean or a within-laboratory standard deviation was marked as a straggler if the outlier test result was significant at the 5% level, and marked as an outlier if the outlier test result was significant at the 1% significance level. Following ISO 5725-2 recommendations, outliers were discarded and stragglers retained unless no other explanations for the outlying observations were found.

Repeatability standard deviation (s_r; within-lab) and reproducibility standard deviation (s_R; between-labs) were used as measurements of precision. The ratio of the repeatability standard deviation and the reproducibility standard deviation (s_R:s_r) of the log-concentration was determined and used as an indicator of the (dis)agreement between the mean results of the laboratories. Ratios up to 3 were considered to represent good agreement, ratios between 3 and 6 to represent fair agreement, and >6 were considered to mean that agreement between the laboratories needed to be improved.

Relative standard deviation (RSD) was used to assess the fluctuations in the data relative to the data mean. Expressed in percentage terms, the formula for RSD is: (sd/mean log concentration)*100. RSD values and associated thresholds represent an attempt to define absolute levels of acceptable sample-to-sample result variability (repeatability, r) and lab-to-lab result variability (reproducibility, R). Standards for RSD have been developed in the literature in an attempt to define absolute levels of acceptable variability in sample-to-sample measurements. Criteria for the analysis were based on Wragg et al. (2011) and Ashley et al. (2012) who suggest that the RSD for reproducibility should be less than

20%, and Wragg et al. who further suggest that RSD for repeatability should be less than 10%.

3. Results

The five laboratories performed bioaccessibility testing on the same six distinct metal-containing materials in four simulated biological fluids. A total of 70 datasets were generated: seven time points with up to ten metal/test substance extractions each. However, some datasets were excluded from analyses as described in Section 3.1.

3.1 Data Exclusion

3.1.1 Quality Control of Protocol Implementation

Differences in protocol implementation between labs identified as part of the quality assurance exercise (see Section 2.5) are summarized in detail in the Supplemental Online Material. The outcome of this exercise led to exclusion of several fluid/time point/lab datasets from statistical analyses when the identified deviations from the SOP had potential to impact the experimental procedures, as discussed below.

- For synthetic perspiration, both datasets (24 and 168 h) for Lab D were excluded from analyses of perspiration data as it reported using a different temperature

during extraction (37°C instead of 30°C).

- Four of the five labs demonstrated lower Pb values for the 168 h time point in perspiration compared to 24 h. The reported lower values could be due to Pb ion complexation and subsequent precipitation. Indeed two labs reported seeing precipitation with a naked eye. This phenomenon is likely to be associated with pH changes. Labs A and E reported a drift in pH up to 7.7-7.9 after 168 h (no information on pH was provided by Lab D; Lab B reported pH around 6.5). While these effects are related to the underlying chemistry of metal ion dominated by complexation with fluid constituents and subsequent precipitation effects, they introduce a greater source of variability to the assays. The results from multiple labs suggest that this combination of fluid composition, time point, and loading is less suitable to assess the repeatability and reproducibility of bio-elution tests for Pb. Thus Pb from leaded brass alloy at 168 h was not included in this evaluation.

- Lab E reported significant evaporation in many of the test vessels containing interstitial fluid at both time points, with some data points not reported at all due to 100% evaporation. Therefore, Lab E data was not included in analyses of interstitial fluid.

- Release of Ni from Ni compound in interstitial fluid at 168 h was less than that at 24 h for Labs B, C, and D; while Lab E only had one triplicate reported due to

evaporation (data already excluded). Labs A and B reported observations of precipitation with Ni compounds in this fluid at this time point and Lab B reported a pH shift upwards of ~1 unit in some cases. While related to the underlying chemistry of metal ion interactions (as described above for Pb) in this particular fluid, these effects introduce a greater source of variability to the assays. The results from multiple labs suggest that this combination of fluid composition, time point, and loading is not suitable to assess the repeatability and reproducibility of bio-elution tests for Ni from Ni compound, therefore data from 168 h were not included in this evaluation.

3.1.2 Limitations imposed by limits of detection

The LODs varied depending upon the metal, fluid, loading and analytical methodology used (e.g., AAS-flame or AAS-GF) and are provided in the Supplemental Online Material. Since one of the goals of this study was to determine reproducibility of measurements between labs, the variable LODs precluded the possibility of using the measurements that were below the LOD (only the case for the Inconel alloy), either by substituting them with the LOD or replacing them by a fraction of the LOD. Therefore, all measurements <LOD were noted as such and excluded from any statistical analyses.

Datasets with 2 or more labs reporting results <LOD and therefore excluded from the inter-laboratory validation were only an issue for the release of Fe and Cr from the Inconel alloy; Cr in gastric fluid; Cr and Fe in 24 h perspiration; Fe in 168 h perspiration;

Cr and Fe in 24 and 168 h interstitial fluid; and Cr in 24 h lysosomal fluid.

3.1.3 Precision measures and outliers

As illustrated in Table 3, there were a total of 11 outliers identified among all treatments, with at least one outlier present within each treatment except the 168 h extraction of interstitial fluid. Per ISO 5725-2 recommendations, all outliers were discarded from the database prior to subsequent analyses. Retained datasets (number of labs and number of measurements) are summarized in Table 4.

3.2 *Results from Statistical Analyses*

3.2.1 Repeatability and reproducibility results

For the retained test substances and treatment fluid conditions, the means and measures of repeatability (s_r) and reproducibility (s_R) of the logarithms of the measurements were calculated and presented under each treatment fluid condition in Table 4. General observations based on intra-laboratory and inter-laboratory measurement variability for each treatment conditions are presented below according to their respective s_r and s_R calculations.

3.2.1.1 Gastric 2 h

Laboratory data for bioaccessibility after 2 h in synthetic gastric fluid were available for all but the Cr from Inconel alloy (Table 4). In this treatment condition, Ni from Ni compound measurements were the least variable within and across labs, with Pb from leaded brass alloy and Co from Co compound also demonstrating relatively low variability for both measures. Iron from the Inconel alloy, a dataset with the fewest bioaccessibility measures for the gastric fluid treatment, demonstrated some of the highest variability for both measures.

3.2.1.2 Perspiration – 24 h

For the bioaccessibility dataset after 24 h in synthetic perspiration fluid, data were retained for all but the Cr and Fe from the Inconel alloy (Table 4). Under these conditions, both Ni-containing test substances and the Cu from Cu concentrate demonstrated a combination of low variability for both the repeatability and reproducibility measures. On the other hand, both Co-containing test substances demonstrated some of the highest variability for both measures under these conditions.

3.2.1.3 Perspiration – 168 h

For the extended 168 h exposure to perspiration fluid, the bioaccessibility data were retained for all but the Fe from Inconel alloy and Pb from leaded brass alloy (Table 4). Again, Ni from Ni compound demonstrated relatively little variability within and between labs, along with Zn from leaded brass and Cu from Cu concentrate. Similar to

the 24 h perspiration treatment, Co from Co powder had a relatively high variability for both measures.

3.2.1.4 Lysosomal – 24 h

With the exception of Cr from the Inconel alloy, bioaccessibility measurement data were retained for all metal/test substance analyses in lysosomal fluid for 24 h (Table 4). The measurement variability within and between labs was relatively low for both Ni-containing test substances, Pb from leaded brass alloy, and the Co from Co compound. In contrast, Co from Co powder and Cu from leaded brass alloy had relatively large s_r and s_R values.

3.2.1.5 Lysosomal – 168 h

Bioaccessibility measurement data were retained for all metal/test substance analyses conducted over the extended 168 h period in lysosomal fluid (Table 4). Under these conditions, the variability in measurements both within and between labs was relatively low for Ni from Ni compound, Cr from Inconel alloy, and Co from Co powder. On the other hand, Fe from the Inconel alloy and Cu from Cu concentrate measurements demonstrated relatively high variability for both measures under these conditions.

3.2.1.6 Interstitial – 24 h

For the bioaccessibility dataset after 24 h in interstitial fluids, data that passed QA check and outlier evaluations were available for all but the Cr and Fe measurements from the Inconel alloy (Table 4). In general, the dataset for this treatment condition was the most variable as it relates to both repeatability and reproducibility. Only Ni from Ni compound had relatively low variability for both parameters, whereas the three metals measured from the leaded brass alloy sample (Cu, Pb, and Zn) demonstrated some of the highest variability in the overall dataset.

3.2.1.7 Interstitial – 168 h

For the extended 168 h exposure to interstitial fluid, the bioaccessibility data were not retained for four of the 10 metal/test substance analyses, including Cr and Fe measurements from Inconel alloy, as well as Pb from leaded brass alloy and Ni from Ni compound (Table 4). The measurement variability within and between labs was relatively low for Ni from Inconel alloy and the Co from Co powder. In contrast, Zn from leaded brass alloy had relatively large s_r and s_R values.

3.3 s_R:s_r *ratio results*

As demonstrated in Table 4, the average repeatability standard deviation (s_r) of the log-concentration among all treatment conditions varied slightly (between 0.014 and 0.083), with the exception of interstitial fluid at the 24 h extraction time period. These findings demonstrate good within-lab agreement. However, the between-lab agreement relative to

392 the within-lab agreement was not as satisfactory. This can be illustrated for many of the
393 treatment condition datasets by calculating the ratio of the reproducibility standard
394 deviation (s_R) and the repeatability standard deviation (s_r) of the log-concentration, which
395 was used as an indicator of the agreement/disagreement between the mean results of the
396 laboratories (Table 5). Even after exclusion of measurements obtained outside the SOP
397 (Section 3.1.1) or datasets with more than 2 values below the LOD (Section 3.1.2), the
398 reproducibility standard deviations of log-concentrations for perspiration fluid (24 h and
399 168 h extraction time) and lysosomal fluid (168 h extraction time) remain very large as
400 compared with the repeatability standard deviations. This is reflected in the high s_R:s_r
401 ratios in several of the metals measurements for these treatment conditions. Based on the
402 criteria used to interpret the s_R:s_r ratio the perspiration treatment conditions were poorly
403 reproduced between labs. This is especially true at 24 h for Co from Co compound (24.0)
404 and Co powder (12.7), and all three metals (Cu, Pb, and Zn) measured from leaded brass
405 alloy (19.9, 6.6, and 19.0, respectively). There was fair agreement in variability between
406 repeatability and reproducibility measurements under the gastric and long-term lysosomal
407 treatments (average s_R:s_r for all 10 metal/test substance analyses equal to 3.4 and 5.3,
408 respectively), while the average s_R:s_r ratios for interstitial fluids (24 h and 168 h) and the
409 short-term lysosomal treatment indicated good agreement in variability within and
410 between labs (average s_R:s_r for all 10 metal/test substance analyses equal to 2.2, 2.3, and
411 2.5, respectively).
412
413 From the perspective of the metal/test substance analyses, both Ni-containing substances,
414 the three metals from the leaded brass sample (Cu, Zn, Pb), and Cu from Cu concentrate

all displayed fair inter-laboratory agreement (relative to intra-laboratory agreement) across treatment conditions. The remaining metal/test substance (Fe, Cr) analyses showed poor agreement between repeatability and reproducibility, indicating that the agreement between the laboratories needs to be improved.

3.4 RSD results

Relative standard deviation (RSD) analysis of the log concentration is another way to consider intra- and inter-laboratory measurement variability. This approach examines s_r and s_R measures individually, assessing the fluctuations in the data relative to the log mean. In our study there were only five instances where the standard for repeatability (e.g., 10%) was exceeded (all with metals from the leaded brass sample treated with lysosomal or interstitial fluids) out of a potential 70 treatment+metal/test substance analyses combinations. Figure 1 demonstrates that with the exception of interstitial, all other fluids have fairly low within-lab variability for the time point shown (<4%). This suggests that measurements were satisfactory based on within-lab variability for all treatment conditions (Table 6).

According to the RSD analysis, the inter-laboratory variability appears to be unacceptable (e.g., >20%) in the interstitial fluid treatment (24 h) for Pb and Zn from the leaded brass alloy, and Ni from the Inconel alloy. Additionally, the RSD analysis indicates very large reproducibility RSD values for Co from Co compound (perspiration, 24 h), Cr from Inconel alloy (perspiration, 168 h), Cu in leaded brass alloy (lysosomal, 24 h), and Zn in

leaded brass alloy (interstitial, 24). Figure 2 demonstrates the variability observed between laboratories.

4. Discussion

Bioelution methods have been used extensively as an alternative to *in vivo* testing for evaluation of metals and metal-containing materials over the last 15 years. Existing publications include those evaluating the bioaccessibility of various metals (Co, Ni, Cr, Pb, Zn, Cu, Cd, arsenic, beryllium, manganese, tin, and uranium) from metal compounds, alloys, soils, household dust, welding fumes, and mine waste in various synthetic fluids (Stopford et al., 2003; Stefaniak et al., 2014; Hillwalker and Anderson, 2014; Oller et al., 2009; Hamel et al., 1998; Vasiluk et al., 2011; Drexler and Brattin, 2007; Wragg et al., 2011; Ellickson et al., 2001; Turner, 2011; Gray et al., 2010; and Twining et al., 2005; Mazinanian et al., 2013; Hedberg et al., 2013). A series of studies published by the KTH laboratory, primarily reported on the bioaccessibility of Fe, Cr, and Ni from various alloys and metals (Herting et al., 2008; Hedberg et al., 2010; Mazinanian et al., 2013; Midander et al., 2010; Hedberg and Odnevall Wallinder, 2013; Hedberg et al., 2013; Jiang et al., 2012; Hedberg et al., 2011; Stockmann-Juvala et al., 2013).

In recent years, various metals associations have also used bioaccessibility methods to meet regulatory requirements imposed under REACH. Prior to REACH, precedents for the use of bioaccessibility in regulatory frameworks already existed. For example, the

European standard for release of nickel in artificial perspiration (BS EN 1811, 2011) has also been incorporated into Europe's Classification, Labelling And Packaging of Substances and Mixtures Regulation (CLP); this regulation stipulates that Ni-containing alloys be classified according to the amount of nickel released using this method (EC, 2008). Another example is the restriction of 19 metals in consumer articles that can be mouthed by children based on the use of EN71.3 (EC, 2013). In the United States, the soluble (bioaccessible) cadmium in surface coatings of children's jewelry is also restricted (US CPSC, 2008; ASTM, F963).

As evidenced by the number of recent publications on this topic, a variety of fluid compositions and protocols for performing bioaccessibility testing exist. While these are generally similar in nature, it was important in this inter-laboratory study to establish one SOP that could be followed by each of the participating laboratories. The methods and simulated fluids were selected based on their relevance to oral, inhalation and dermal exposure; those previously published by Stopford et al. (2003) served as the basis of developing the SOP.

With regards to gastric fluid, the protocol of ASTM D5517 (2007) was employed for the estimation of metal solubility in the stomach. Synthetic gastric fluid extractions such as this one have been compared with the *in vivo* solubility of lead silicates in the stomach of rats (Ruby et al., 1999) and more recently with the acute oral toxicity in rats exposed to nickel compounds (Henderson et al., 2012). While additional compartments such as saliva and intestinal fluids can be informative in assessing the bioavailaiblity of some

484 metals, these fluids were not included in the present validation program. The ASTM
485 D5517 (2007) protocol was also followed for extractions with simulated interstitial and
486 lysosomal fluids; with the interstitial fluid closely matching Gamble's solution. The
487 interstitial fluid represents lung fluid and uses citrate in place of proteins while acetate is
488 used to represent organic acids. The interstitial fluid has been used to compare the
489 pulmonary durability of inhaled man-made fibers (Ziotos et al., 1997; Leheude et al.,
490 1997). The solubility of substances that have been phagocytized and subsequently
491 released into the intracellular environment has been estimated using lysosomal fluid (de
492 Meringo et al., 1994; Theolahn at al., 1994). This fluid includes glycine, a variety of salts
493 of organic acids, and citric acid. Citric acid and other organic acids in lysosomal fluid are
494 known to form complexes with metals, resulting in increased release of metals (Hedberg
495 et al., 2010; Hedberg et al., 2011; Hillwalker and Anderson, 2014). Finally, the synthetic
496 perspiration fluid cited in standard EN 1811 (2011) and approved by the European
497 Committee for Standardization (CEN) in 1998 was used here to simulate the release of
498 soluble metal onto skin. Other compositions for artificial perspiration have also been
499 tested (e.g., Stefaniak et al., 2014). Hillwalker and Anderson (2014) compared the
500 bioaccessibility results from a variety of alloys (Stainless steels AISI 304 and 316,
501 Inconel, Monel) in fluids with slightly different compositions and concluded that Ni and
502 Cr absolute releases from alloys are especially sensitive to fluid composition and
503 extraction time.
504
505 In the current study, analyses of repeatability measures using two different approaches
506 (s_R:s_r ratios and RSD) show that the within-laboratory variability was generally

satisfactory for all treatment conditions with the exception of some metals in interstitial fluid (Tables 5 and 6). However, variability between laboratories was found to exceed accepted criteria, the extent of which depended on whether the s_R:s_r ratios or the RSD approaches were used. Using the ratio of s_R:s_r, the inter-laboratory concordance for synthetic perspiration was found to be poor overall (ratios >6; see Table 5). Testing in gastric and 168h lysosomal fluids resulted in fair agreement between labs (ratios = 3-6), while testing in interstitial and 24h lysosomal fluids resulted in good agreement in variability within labs (ratios <3). Similarly, while RSD analysis showed better agreement between laboratories overall, higher inter-laboratory than within-laboratory variability was observed.

A study aimed at evaluating analytical procedures among labs was conducted prior to initiating the present round robin bioaccessibility study. Samples of interstitial fluid spiked with known metal concentrations (blank, Co, Cu, Ni, Pb, and Zn) were provided (in blind fashion) to each of the laboratories to determine the analytical concentrations. After eliminating outliers, the statistical analysis resulted in an s_R:s_r ratio of about 6, indicating a lack of harmonization among laboratories (data not shown). As a result of this analytical exercise, several recommendations for improving reproducibility were subsequently implemented in the SOP utilized in the bioaccessibility inter-laboratory exercise.

Still, careful comparison of each of the laboratory reports for the round robin revealed that the SOP might not have been precise enough for some parameters (e.g., buffering

method). A systematic comparison between the SOP and the reports from the 5 labs also identified a number of methodological differences. For interstitial fluid, the method of CO_2 buffering varied widely among all 5 labs including equipment, location (headspace, fluid, or chamber), and moisturizing gas, etc. Although this is a potential major source of variation, and even though all labs performed this step differently, no clear association between the results for this fluid and any specific lab was identified. Another difference observed between labs was the incidence of evaporation in some fluids. Lab E reported evaporation at 24h in interstitial fluid while Labs A, B, C, and E reported evaporation over time and difficulty measuring/maintaining pH in this fluid. Also in interstitial fluid, Lab A noted precipitation with Ni compound and Pb from leaded brass alloy and Lab B reported precipitation with Ni compound. This precipitation may have been due in part to the evaporation taking place in the vessels. Control of pH, particularly in the lysosomal fluid, also presented challenges. This issue was also noted in the Unified BARGE Method (UBM) study, which concluded that tighter control of pH was critical in gastric fluid (Wragg et al., 2011). Finally, when measurements approach the limit of determination (e.g., <25 µg/g; but even <100 µg/g), the reproducibility outcomes worsened.

Several lessons can be learned from this exercise. The SOP used in this study had too many degrees of freedom as written, and as such, additional details should be incorporated into future drafts. Substances that are being compared (e.g., Cu metal and Cu alloy) should always be tested side-by-side or at least in the same lab. The choice of particle loading is crucial to minimize effects such as agglomeration and abrasion

553 (Hedberg et al., 2010; Henderson et al., 2012; Stopford et al., 2003; Turner, 2011). On
554 the other hand, it is possible that higher sample loadings could overcome the variability
555 associated with low metal releases close to the LOD. In all cases, realistic conditions
556 need to be considered. It might also be useful to measure metal releases over time (e.g.,
557 μg/g/h) that can better define the kinetics of metal release (Herting et al., 2008; Hedberg
558 et al., 2010; Hillwalker and Anderson, 2014; Stefaniak et al., 2014; Hedberg et al., 2013).
559
560 Limiting longer exposure times when complicating factors such as CO_2 buffering are
561 introduced may reduce inter-laboratory variability. For example, metal complexation and
562 precipitation and difficulties in maintaining the pH may provide an explanation for the
563 change in repeatability observed between 24 and 168 hours in some fluids. In particular,
564 this is an example of why longer time points (168h) may be pushing the limitations of
565 experimental methods where pH, precipitation, changes in volume, buffering, etc. can all
566 introduce variation. Improvements to the SOP are clearly needed to obtain better within
567 and between laboratory agreements. Recommendations for refining the SOP include
568 better defining pH control measures, CO_2 buffering technique, and agitation methods, and
569 ways to minimize evaporation. This is especially true for the interstitial fluid, which
570 stands out as a fluid that requires the most improvement.
571
572 It is useful to compare the results of the current study to those of similar inter-laboratory
573 validation studies of specific bioelution methods. In the study of Drexler and Brattin
574 (2007) an *in vitro* relative bioaccessibility leaching procedure (RBALP) designed to
575 mimic oral Pb exposure conditions was performed by three laboratories on 19 different

576 test materials. The results of each lab were subsequently compared to *in vivo* relative
577 bioavailability (RBA) measures. The authors reported that the intra- and inter-laboratory
578 *in vitro* results were "highly reproducible" with a coefficient of variation (e.g., RSD)
579 equal to 6% and 4%, respectively, and concluded that the RBALP method could reliably
580 estimate Pb RBA *in vivo*. Another round-robin study looked at a different
581 physiologically-based method for estimating the bioaccessibility of Pb, as well as Cd and
582 As, from soils (Wragg et al., 2011). The UBM method, which includes synthetic saliva,
583 gastric and intestinal fluids, was used to assess metal release from As, Cd, and Pb
584 samples. Measurements from seven laboratories were compared to *in vivo* RBA data and
585 the overall outcomes were evaluated based on a set of four benchmark criteria. Results of
586 the UBM method were reported to have met the inter-laboratory criteria for As (RSD =
587 7.43% for stomach phase and 15.72% for stomach + intestine phase). However,
588 compliances for the stomach phase only for Pb (RSD = 22.78%) and stomach plus
589 intestine phases for Cd and Pb (RSD = 35.35% and 81.39%, respectively) were above the
590 benchmark criteria (ie, RSD ≤20%). The authors suggested that tighter control of gastric
591 pH may be helpful and noted that a follow up inter-laboratory study would be needed.
592
593 Using the same RSD criteria the results of the current study appear to be in line with
594 those of Wragg and colleagues (2011), with the possible exception of interstitial fluid at
595 24h (Table 6). In the context of some other studies of similar characteristics it is possible
596 that the criteria used here (RSD ≤10% and ≤20% for intra- and inter-laboratory
597 variability, respectively) may be too stringent. An RSD of 30% or even 40% may be a
598 more realistic cut-off for determining acceptable variation between laboratories. For

example, in one study using a saliva migration test for organic plasticizers, where 15 labs performed validation of the SOP, an RSD of 30% was found to be the best obtainable reproducibility (EUR 19826 EN, 2001). Similarly, in a study to validate a method for environmental assessment of metals, Skeaff et al. (2011) reported that the inter-laboratory variability ranged according to analysis by $CV_{\%}$ (similar to % RSD). In this study, 12/37 measurements had CV% values between 25-56% and 10/37 had values ≥57%. If an RSD of 30% or 40% had been used as the standard for the current study, all between laboratory reproducibility would have been deemed acceptable for all metals and treatment conditions, with the exception of Cr from Inconel alloy in 168h perspiration fluid and Zn from leaded brass alloy in 24h interstitial fluid.

The above discussion applies exclusively to estimates of absolute metal release. However, for most applications, only measures of relative metal release from two or more forms of the same metal are needed, diminishing the requirement for high inter-laboratory reproducibility in absolute metal releases. The high within-laboratory repeatability supports the use of these methods for the assessment of relative metal release and calculation of effective concentration of metals in complex materials where a matrix effects can be present.

In the current exercise we included two alloy samples (Inconel and leaded brass alloys) but we did not include the pure metal components of these alloys (e.g., Cr, Fe, Ni in case of Inconel) as reference materials. Thus effective concentrations of metals in these alloys cannot be calculated based on the data from the present round robin. However, two

622 laboratories that participated in this study previously tested the same sample of a Ni metal
623 powder in lysosomal fluid (Mazinanian et al., 2013; KMHC, 2010). Based on the Ni
624 releases from Ni metal and Inconel alloy in 24h lysosomal fluid, the effective
625 concentration of Ni in Inconel alloy can be calculated as 0.05 and 0.2%, for Mazinanian
626 et al. (2013) and KMHC (2010), respectively (calculations not shown). Using different Ni
627 metal and Inconel samples, an effective concentration of Ni in Inconel of 0.4% was
628 calculated, based on bioaccessibility data in lysosomal fluid at 72 hours reported by
629 Hillwalker and Anderson (2014). In summary, three different laboratories calculated
630 similar effective concentrations of Ni metal in Inconel alloy (relevant to the inhalation
631 route of exposure) even when using different alloys and nickel metal samples and with
632 slightly different absolute releases. The effective concentration of Ni in a SS316 alloy
633 has been recently shown to be a better predictor of *in vivo* inhalation toxicity than its
634 content (Stockmann-Juvala et al., 2013).
635
636 In general, this approach could be applied for the classification of alloys based on
637 classifications of their constituent metals. The relative bioaccessibility in gastric,
638 perspiration and lysosomal fluids could allow the calculation of effective concentration of
639 classified metals in alloys and permit more toxicologically relevant classifications when
640 effective concentrations are compared to classification cut-off limits for mixtures. A
641 similar approach could be applied to other complex materials, such as ores and
642 concentrates, where matrix effects are suspected.
643
644

5. Conclusion

In conclusion, the outcome of this inter-laboratory validation exercise for bioelution testing of metals demonstrates overall satisfactory within-laboratory variability in bioaccessibility data for synthetic gastric fluid, lysosomal fluid, interstitial fluid, and perspiration for all treatment conditions. With regards to between laboratory agreement, a higher inter-laboratory than within-laboratory variability in bioaccessibility results was observed for most metals and treatment conditions suggesting that, for the methods tested, the absolute bioaccessibility results in some biological fluids may not always be in line among different laboratories. There are a number of potential sources of variation that may have contributed to this outcome. The most reproducible results were typically observed with shorter extraction times. The inter-laboratory exercise suggests that the degrees of freedom within the SOP need to be addressed to achieve better concordance in absolute metal releases. However, for hazard and risk assessment applications, the use of these methods to generate relative release data for read-across purposes or to calculate effective concentration of metals in alloys and other complex materials appears to be acceptable.

Acknowledgements

We would like to thank the following individuals for their help with data analysis and/or manuscript preparation: Dr. Wendy Hillwalker (OSU) and Dr. Jon Urban (ToxStrategies).

The project was supported by the following: the Nickel Producers Environmental Research Association, Inc. (Durham, NC, USA), the Cobalt Development Institute (Guildford, Surrey UK), the International Zinc Association (Brussels, Belgium), the European Copper Institute (Brussels, Belgium), Eurometaux (Brussels, Belgium) and the Food Safety and Environmental Stewardship Program at OSU.

References (1430 words)

ASTM (American Society for Testing and Materials), 2007. Standard test method for determining extractability of metals from art materials. In: ASTM, Annual Book of ASTM Standards, vol. 06. 02, D5517–03. ASTM, Philadelphia.

ASTM (American Society for Testing and Materials), 2011. Standard Consumer Safety Specification for Toy Safety. In: ASTM, Annual Book of ASTM Standards, vol. 15. 11, F963-11. ASTM, Philadelphia.

Ashley, K, Shulman, S.A., Brissonb, M.J., Howe, A.M., 2012. Interlaboratory evaluation of trace element determination in workplace air filter samples by inductively coupled plasma mass spectrometry. J. Environ. Monit. 14(2), 360-367.

Brock, T., and Stopford, W., 2003. Bioaccessibility of metals in human health risk assessment. Evaluation of risk from exposure to cobalt compounds. J Env. Monit. 5, 71N-76N.

BS (British Standard) EN 1811, 2011. Reference test method for release of nickel from all post assemblies which are inserted into pierced parts of the human body and articles intended to come into direct and prolonged contact with the skin.

BS (British Standard) EN 71-3, 2013. Safety of toys. Migration of certain elements

Cordeiro, F., Baer, I., Robouch, P., Emteborg, H., Got, J.C., Kortsen, B., de la Calle, B., 2012. IMEP-34: Heavy metals in toys according to EN 71-3:1994; Interlaboratory Comparison Report. JRC Scientific and Policy Reports. EUR 25380 EN.

de Meringo, A., Morscheidt, C., Thélohan, S., Tiesler, H., 1994. *In vitro* assessment of biodurability: Acellular systems. Environ. Health Perspectives. 102 Suppl 5, 47-53.

Drexler, J.W., Brattin, W.J., 2007. An in vitro procedure for estimation of lead relative bioavailability: with validation. Hum. Ecol. Risk Assess. 13, 383–401.

EC (European Council), 2006. Regulation (EC) No. 1907/2006 of the European Parliament and of the Council of 18 December 2006 Concerning the Registration, Evaluation, Authorisation and Restriction of Chemicals (REACH). Official Journal of the European Union L396: L136/133-L136/280.

EC (European Commission), 2008. Regulation (EC) No. 1272/2008 on Classification, Labelling and Packaging (CLP) of Substances and Mixtures.

EC (European Commission), 2013. Toy Safety Directive 2009/48/EC.

ECHA (European Chemicals Agency), 2008. Guidance on Information Requirements and Chemical Safety Assessment. Chapter R. 6: QSARs and grouping of chemicals. ECHA, Helsinki.

ECHA (European Chemicals Agency), 2013. Guidance on the Application of the CLP Criteria. version 4.0 ECHA, Helsinki.

ECHA (European Chemicals Agency), 2014. Available from: http://echa.europa.eu/view-article/-/journal_content/title/rac-and-seac-agree-on-restrictions-and-authorisations. Last accessed April 9, 2014.

Ellickson, K.M., Meeker, R.J., Gallo, M.A., Buckley, B.T., Lioy, P.J., 2001. Oral

bioavailability of lead and arsenic from a NIST standard reference soil material. Arch. Environ. Contam. Toxicol. 40, 128–135.

EUR 19826 EN, 2001. Validation of methodologies for the release of di-isonoylphthalate (DINP) in saliva simulant from toys.

Gray, J.E., Plumlee, G.S., Morman, S.A., Higueras, P.L., Crock, J.G., Lowers, H.A., Witten, M.L., 2010. In vitro studies evaluating leaching of mercury from mine waste calcine using simulated human body fluids. Environ. Sci. Technol. 44, 4782–4788.

Guney, M., Zagury, G.J., 2014. Bioaccessibility of As, Cd, Cu, Ni, Pb, and Sb in Toys and Low-Cost Jewelry. Environ. Sci. Technol. 48, 1238−1246.

Hamel, S.C., Buckley, B., Lioy, P.J., 1998. Bioaccessibility of metals in soils for different liquid to solid ratios in synthetic gastric fluid. Environ. Sci. Technol. 32, 358–362.

Hedberg, Y., Gustafsson, J., Karlsson, H.L., Möller, L., Odnevall Wallinder, I., 2010. Bioaccessibility, bioavailability and toxicity of commercially relevant iron- and chromium-based particles: in vitro studies with an inhalation perspective. Part. Fibre Toxicol. 7, 23.

Hedberg, Y., Hedberg, J., Liu, Y., Odnevall Wallinder, I., 2011. Complexation – and

ligand-induced metal release from 316L particles – importance of particle size and crystallographic structure. Biometals 24, 1099–1114.

Hedberg, Y., Hedberg, J., Odnevall Wallinder, I., 2012. Particle characteristics and metal release from natural rutile (TiO_2) and zircon particles in synthetic body fluids. J. Biomater. Nanobiotechnol. 3, 37-49.

Hedberg, Y., Mazinanian, N., Odnevall Wallinder, I., 2013. Metal release from stainless steel powders and massive sheet – comparisons and implications for risk assessment of alloys. Environ. Sci. Processes Impacts, 65: 135-146

Hedberg, Y., Odnevall Wallinder, I., 2013. Metal release and speciation of released chromium from a biomedical CoCrMo alloy into simulated physiologically relevant solutions. J. Biomed. Mater. Res. B Appl. Biomater. 102(4), 651–895.

Henderson, R.G., Cappellini, D., Seilkop, S.K., Bates, H.K., Oller, A.R., 2012. Oral Bioaccessibility Testing and Read-Across Hazard Assessment of Nickel Compounds. Regul. Toxicol. Pharmacol. 63(1), 20-28.

Herting, G., Wallinder, I.O., Leygraf, C., 2008. Metal release rate from AISI 316L stainless steel and pure Fe, Cr and Ni into a synthetic biological medium-a comparison. J. Environ. Monit. 10, 1092–1098.

Hillwalker, W.E., Anderson, K.A., 2014. Bioaccessibility of metals in alloys: Evaluation of three surrogate biofluids. Environ. Pollut. 185, 52-8.

Jiang, T., Odnevall Wallinder, I., Herting, G., 2012. Chemical stability of chromium carbide and chromium nitride powders compared with chromium metal in synthetic biological solutions. ISRN Corrosion 2012, Article ID 379697.

KMHC (Kirby Memorial Health Center), 2010. Kirby Memorial Health Center Compiled Analysis Reports to NiPERA, Inc. for 15 Nickel Substances: Solubility in Simulated Fluids.

Lehuede, P., de Meringo, A., Bernstein, D.M., 1997. Comparison of the chemical evolution of MMVF following inhalation exposure in rats and acellular *in vitro* dissolution. Inhal. Toxicol. 9(6), 495-523.

Mazinanian, N., Hedberg, Y., Wallinder, I.O., 2013. Nickel release and surface characteristics of fine powders of nickel metal and nickel oxide in media of relevance for inhalation and dermal contact. Regul. Toxicol. Pharmacol. 65, 135–146.

Midander, K., Pan, J., Leygraf, C., 2006. Elaboration of a test method for the study of metal release from stainless steel particles in artificial biological media. Corrosion Science 48(9), 2855–2866.

Midander, K., de Frutos, A., Hedberg, Y., Darrie, G., Wallinder, I.O., 2010. Bioaccessibility studies of ferro-chromium alloy particles for a simulated inhalation scenario: a comparative study with the pure metals and stainless steel. Integr. Environ. Assess. Manag. 6, 441–455.

ISO (International Organization for Standardization) 5725-2, 1994. Accuracy (trueness and precision) of measurement methods and results, Part 2: Basic method for the determination of repeatability and reproducibility of a standard measurement method. International Organization for Standardization, Geneva.

Oller, A.R., Cappellini, D., Henderson, R.G., Bates, H.K., 2009. Comparison of nickel release in solutions used for the identification of water-soluble nickel exposures and in synthetic lung fluids. J. Environ. Monit. 11, 823–829.

Ruby, M.V., Schoof, R., Brattin, W., Goldade, M., Post, G., Harnois, M., Mosby, D.E., Casteel, S.W., Berti, W., Carpenter, M., Edwards, D., Cragin, D., Chappell, W., 1999. Advances in evaluating the oral bioavailability of inorganics in soil for use in human health risk assessment. Environ. Sci. Technol. 33, 3697–3705.

Skeaff, J.., Adams, W.J., Rodriquez, P., Brouwers, T., Waeterschoot, H., 2011. Advances in Metals Classification Under the United Nations Globally Harmonized System of Classification and Labeling. Integrated Environ. Assess. Manag. 7(4), 559–576.

Stefaniak, A.B., Duling, M.G., Geer, L., Virji, M.A., 2014. Dissolution of the metal sensitizers Ni, Be, Cr in artificial sweat to improve estimates of dermal bioaccessibility. Environ. Sci. Process Impacts 16(2), 341-51.

Stockmann-Juvala H, Hedberg Y, Dhinsa NK, Griffiths DR, Brooks PN, Zitting A, Wallinder IO, Santonen T., 2013. Inhalation toxicity of 316L stainless steel powder in relation to bioaccessibility. Hum. Exp. Toxicol. 32(11), 1137-1154.

Stopford, W., Turner, J., Cappellini, D., Brock, T., 2003. Bioaccessibility testing of cobalt compounds. J. Environ. Monit. 5, 675–680.

Tomlin, J., Brown, N., Ellis, A., Carlsson, A., Bogentoft, C., Read, N.W., 1993. The effect of liquid fibre on gastric emptying in the rat and humans and the distribution of small intestinal contents in the rat. Gut 34, 1177–1181.

Turner, A., 2011. Oral bioaccessibility of trace metals in household dust: a review. Environ. Geochem. Health 33, 331–341.

Twining, J., McGlinn, P., Loi, E., Smith, K., Gieré, R., 2005. Risk ranking of bioaccessible metals from fly ash dissolved in simulated lung and gut fluids. Environ. Sci. Technol. 39, 7749–7756.

US CPSC (United States Consumer Product Commission), 2008. Consumer Product Safety Improvement Act, Public Law 110-314.

Vasiluk, L., Dutton, M.D., Hale, B., 2011. In vitro estimates of bioaccessible nickel in field-contaminated soils, and comparison with in vivo measurement of bioavailability and identification of mineralogy. Sci. Total Environ. 409, 2700–2706.

Wang, S.W., Lu, K.Y., Chen, S.M., Young, T.K., 2001. Gastric emptying and intestinal transit of liquid and solid markers in rats with chronic uremia. Chin. J. Physiol. 44, 81–87.

Wragg, J., Cave, M., Basta, N., Brandon, E., Casteel, S., Denys, S., Gron, C., Oomen, A., Reimer, K., Tack, K., Van de Wiele, T., 2011. An inter-laboratory trial of the unified BARGE bioaccessibility method for arsenic, cadmium and lead in soil. Sci. Total Environ. 409, 4016–4030.

Zoitos, B.K., de Meringo, A., Rouyer, E., Thelohan, S., Bauer, .J, Law, B., Boymel, P.M., Olson, J.R., Christensen, V.R., Guldberg, M., Koenig, A.R., Perander, M., 1997. In vitro measurement of fiber dissolution rate relevant to biopersistence at neutral pH. An interlaboratory round-robin. Inhal. Toxicol. 9(6), 525–540.

Figures – each 1.5 to 2-column fitting images

Figure 1. Within-laboratory variability. All fluids except interstitial fluid have fairly low within-lab variability (<4%) for the time point shown (2 h, gastric; 24 h, all others). %RSD = percent relative standard deviation.

Figure 2. Between-laboratory variability. Results varied between laboratories depending on the metal and fluid tested. As shown here, gastric and lysosomal fluids had more reproducibility than other fluids at the time point shown (2 h, gastric; 24 h, all others). %RSD = percent relative standard deviation.