



Rapid Market Screening Program

Annex B

Quality Assurance and Quality Control

Quality Assurance and Quality Control

For this report, “quality assurance” refers to measures taken to encourage high quality data collection and analysis, such as the development of, and training in a standard RMS implementation protocol (available as Annex A). “Quality control” refers to processes to identify and correct errors, omissions, or other inaccuracies in the RMS process or data, including confirmatory testing of a subset of samples to validate or invalidate field XRF readings.

Quality Assurance

Pure Earth’s primary means of encouraging quality data collection and analysis included the development of a standardized RMS Protocol guiding the methods and sequencing of each step of the program, the training of RMS Investigator teams in that protocol, and ongoing supervision and re-training of teams as the program was implemented.

Pure Earth divided the implementation of the RMS program into three phases: a Formative Research Phase in which we piloted the methodology in several countries to troubleshoot any implementation hurdles; Phase 1, in which the first 15 countries implemented the program; and Phase 2, in which the final 10 countries implemented the program. After each phase, the RMS team reviewed challenges and lessons from the implementation and made any necessary revisions to the protocol and training. Notable changes made after the formative phase included: guidance to specify and narrow the types of items purchased and analyzed; additional guidance on the number of samples desired; and clarification of the XRF setting to use during analysis. After Phase 1, the protocol was amended to shift from analyzing wet paint samples to dried paint swatches. Aside from the paint analysis method, the protocol did not differ substantially between phases and the analytical results from Phases 1 and 2 are included here.

Quality Control

The primary quality control measures in the RMS program included:

1. Regular communication with, and oversight of RMS Investigator teams to ensure they understood the protocol, felt comfortable and confident in their plans, and to address any challenges or questions that arose.
2. Ongoing reviews of data uploaded into the SurveyCTO database platform to ensure that RMS Investigators collected appropriate sample types and numbers of samples, and that sample logs contained all required information.
3. A review of descriptions and item categorization choices made by RMS Investigators (e.g., are leafy herbs a spice, a medicine, or “other foods?”) to ensure consistency with the RMS protocol and the descriptions and categorization choices made by other country teams
4. A review of XRF field data uploaded by RMS Investigator teams to the SurveyCTO database to identify any formatting, unit, or input errors and to highlight any outliers that might require further quality control inquiries.

5. Shipping of a subset of more than 1,000 samples to New York for confirmatory analysis with an XRF analyzer in Pure Earth's headquarters that was operated by an experienced expert, is known to be in good working order, has shown consistency with lab results, and was regularly calibrated against a "standard" sample with a known lead concentration
6. Confirmatory analysis of a subset of 354 samples by accredited laboratories using analytical methods known to be more accurate and sensitive than XRFs

Summary of XRF Performance as a Screening Tool

Lead concentrations measured with portable XRF analyzers proved highly comparable to results from laboratory analysis with the following two exceptions. First, the XRFs appeared to inflate lead concentrations among four samples for which lab results exceeded 30,000 ppm (3% lead), suggesting that the XRF's accuracy may have decreased at very high lead concentrations. Second, confirmatory screening with a New York-based XRF and lab testing suggested that the XRF used in both Tajikistan and Kazakhstan provided anomalous field results; accordingly the field data from these countries were expunged by the RMS Quality Control team. We do not know if this was due to equipment malfunction, contamination of the XRF measurement window, or another type of user or mechanical error. In the case of these two countries, only samples that were available for analysis by a different XRF were included in this report. Finally, we also note that we could not compare XRF results with lab results for items with comparatively low lead concentrations due to the lower detection limit of the XRFs, which for most materials is approximately 2-4 ppm, compared to 0.2 ppm to 0.5 ppm in the lab. This fact does not suggest inaccuracy among the XRFs but is simply an analytical limitation of the device.

Overall, the XRFs proved to be an excellent screening tool for detecting elevated lead levels in products, particularly among items for which the reference value exceeds the XRF's lower detection limit. More details regarding correlations between XRF and lab results are provided below.

Quality Control Methodology

Each team of RMS Investigators collected and analyzed samples with an XRF following the RMS Protocol. Sample information was entered into Survey CTO and reviewed by the RMS Quality Control team. Country teams selected samples to ship to New York for potential laboratory testing according to RMS Protocol, as follows:

- Foodware (metallic, ceramic and plastic pots, pans, cups, bowls, and serving implements): From each city, select 2 of each item with detectable lead; pots and pans are the priority, e.g., 6 aluminum cooking pots (priority), 6 ceramic cooking pots (priority), cups or bowls, and 3 plastic cups or bowls. If no lead is detected, select 3 items from each city, one metal, one ceramic, and one plastic.
- Cosmetics: XRF testing only is anticipated if XRF detection is at least 10 ppm, which is the reference level. If samples are collected for laboratory analysis, from each city, select 2-3 samples of lipstick and/or products used on children.

- Toys: From each city, select 2 toys (include plastic and painted) with detectable lead. If no lead is detected, select 2 plastic baby toys and 1 painted toy.
- Spices: From each city, select 3 to 5 samples with detectable lead. If no lead is detected, select one sample of each type from each city. Staple Dry Foods: All samples (~3 from each city) should be sent to New York for laboratory testing because reference values are lower than XRF's lower detection limit.

Samples that were shipped were itemized on country-specific spreadsheets that included: sample ID numbers, descriptions, dates of collection, field XRF results, and any other pertinent information, with such information downloaded from SurveyCTO and verified by the RMS Team.

More than 1,000 samples (approximately 20% of all samples collected and an average of about 40 samples per country) were shipped to New York for confirmatory analysis by a second XRF, an accredited lab, or both. The samples received were compared with the country-specific spreadsheet of items shipped, ensuring that each item was present and appropriately labeled. Approximately 1,000 additional XRF measurements of selected items were made in New York to confirm XRF readings reported by country teams, obtain a value for non-detect readings when not reported, further evaluate XRF readings in heterogeneous samples such as toys, and to evaluate the effect of metal components (e.g., wires, batteries, lights, gear boxes) on XRF readings in certain items such as toys. These additional XRF measurements followed the same analytical methods outlined in the RMS Protocol.

Comparisons Between Field XRF And New York-Based XRF Results

The XRF readings from country teams were compared to confirmatory XRF readings of a subset of samples shipped to New York. Field XRF readings that differed by more than 30% from New York-based XRF readings were flagged for further evaluation. In general, XRF results from the field compared favorably to those measured in New York, with most discrepancies relating to sample heterogeneity (e.g., different parts of a toy measured in the field versus in New York) or field readings impacted by internal metal parts (e.g., wires or batteries inside toys). A detailed account of field XRF versus New York-based XRF results will be available as Annex B on the RMS page of Pure Earth's website.

Expunging Field Data From Certain Products From Tajikistan And Kazakhstan

Despite the general consistency between field XRF, New York-based XRF, and lab results, the field XRF readings from Tajikistan and Kazakhstan had significant discrepancies across several item types when compared to New York-based XRF and lab results. Ultimately, all field data from these countries were expunged. Only samples that were available for analysis by a New York-based XRF were included.

A variety of samples shipped from Tajikistan and Kazakhstan showed field XRF readings in the tens to hundreds of ppm lead, but showed "non-detect" levels according

to the New York-based XRF (i.e., <3 or 4 ppm) and lab results (i.e., <0.50 ppm). These discrepancies could not be accounted for by sample heterogeneity, interference by internal metal parts, or other unique characteristics of the samples. The discrepancies resulted in the Quality Control Team's decision to expunge the field datasets and replace them with a more limited set of New York-based XRF screening results.

Importantly, the same XRF was used to screen samples from both Tajikistan and Kazakhstan, and this XRF was not used in any other countries. It appears that the XRF may have been corrupted at some point during the screening. In consideration of these discrepancies, the Quality Control Team concluded that the field XRF data from Tajikistan and Kazakhstan should not be reported. New York-based XRF data is reported when results from 3 or more samples were available. Note that samples collected in Tajikistan and Kazakhstan that were not shipped to New York were discarded before additional XRF or lab testing could be completed.

Selection Of Samples And Procedure For Laboratory Testing

A total of 356 samples were subjected to lab analysis. The Quality Control Team selected samples representing each of the item categories as well as subcategories within each category (e.g., among cosmetics, a selection of lipsticks, eyeliners, face powders). Priority was given to items intended for children. Items representing a range of XRF results were selected in each category. Samples of staple dry foods and spices were submitted for lab testing regardless of whether they had non-detect XRF readings, as the XRF's lower detection limit generally exceeded the reference value for such products. A summary of the samples submitted for lab testing is included in the table below.

Item Type	Number of Samples Submitted for Laboratory Testing for Total Lead	Comments
Ceramics	1	pottery shard
Cosmetics	64	lipstick, nail polish
Staple Dry Foods	65	flours, rice, etc.
Miscellaneous	9	medicine, sleeping mat, ayurvedics
Other Foods	15	sweets, miscellaneous
Plastic Foodware	17	
Spices	125	

Toy - Painted	3	
Toy - Plastic/Rubber	57	
Total =	356	

Most samples were submitted to Complete Environmental Testing, Inc. in Stratford, Connecticut. The samples were microwave extracted with acid using Method 3051, and the extract was analyzed by inductively coupled plasma mass spectrometry (ICP-MS) by EPA Method 6020A. Detection limits generally ranged from 0.20 to 0.50 mg/kg (equivalent to ppm). The lab reports contain the associated Chain of Custodies for the samples. Selected perishable samples (spices, staple dry foods) from Kyrgyzstan, Pakistan, Türkiye, Tunisia, Azerbaijan, Armenia, Georgia, Uganda, and Egypt were submitted for testing at the MULTITEST laboratory in Georgia using method GOST 30538-97, an atomic absorption spectroscopy method for metals in food with a minimum detection limit of 0.10 mg/kg. Selected perishable samples from Indonesia and the Philippines were submitted to accredited F.A.S.T. Laboratories in the Philippines. The laboratory analyzed the samples by flame atomic absorption spectroscopy with a minimum detection limit of 0.10 mg/kg according to the *Official Method of Analysis of AOAC International, 20th ed. 2016*.

Comparison between XRF and Lab Results

A total of 356 samples were submitted for lab testing to confirm lead concentrations measured by XRF, as summarized below. Because the RMS protocol dictated taking multiple XRF readings for all samples, a representative XRF result was selected for comparison against the lab results. In general, where multiple XRF readings were present (i.e., from the country teams and/or with additional XRF readings completed in New York), an average of all readings was used with the following exceptions:

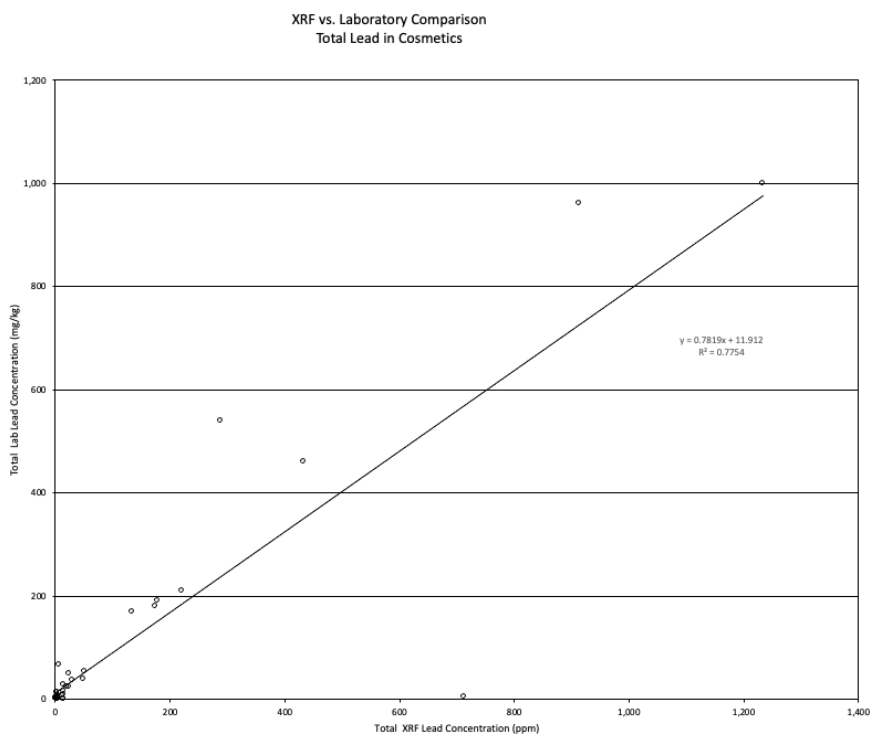
- XRF readings were available for the portion of a heterogeneous sample that was submitted to the laboratory (e.g., the red portion of a toy). In this case, the XRF reading(s) for that portion were used for comparison purposes.
- XRF readings completed by country teams were adversely impacted by the presence of metal interior to the toy upon inspection of those items in New York, as supported by additional XRF readings. In this case, the readings by the country teams were omitted from the average.
- XRF readings for particular countries (e.g., Tajikistan and Kazakhstan) were omitted when they were not supported by additional XRF measurements in New York or lab analysis.
- When “non-detect” (ND) values were reported by the XRF, a value of ND<#/ sq. rt. of 2 was used to provide a numeric value for a ND reading, as indicated in the table. When ND values were not reported by the country teams and additional XRF measurements could not be completed (e.g., the samples were not shipped to New York and were tested by a local laboratory), an entry of ND<3 ppm was

used as an average detection limit for most items, as indicated in the table as “presumed ND<3 ppm”.

A summary of the comparison of lead concentrations measured by XRF with those from lab testing is provided below by item category. Country-specific comparisons were not warranted for the lab versus XRF data comparison, but country-specific observations are included herein by category when appropriate. Note that ceramics samples were not submitted for laboratory testing for comparison to XRF, with one exception, for reasons described in the ceramics subsection along with observations relating to XRF measurements in ceramics. Note that wet or dry paint samples that were collected and screened as part of RMS were not shipped to New York and were not otherwise the subject of the comparison with laboratory data.

Cosmetics

A total of 64 cosmetics samples were submitted for laboratory testing, including eye shadow, eyeliner, lipstick, face powders, and henna. Lead was not detected by the laboratory in 14 of the samples. Lead was detected by the laboratory at concentrations below the 2 ppm threshold in 12 of the 64 samples, and at concentrations between 2 ppm and a high of 320,000 ppm in 32 of the 64 samples. The 43 cosmetic samples in which the lab detected lead below 1,000 ppm correlated well with the field XRF results ($R^2 = 0.7819$), as shown in the figure below. In samples where the lab did not detect lead (generally ND <0.50 ppm), the field XRF also generally did not detect lead (ND<3 or ND<4 ppm). Conversely, there were 15 samples in which lead was not detected by XRF (with various minimum levels of detection) where laboratory lead was detected from 1 to 14 ppm. There were two exceptions wherein the field XRF detected lead at 4 and 12 ppm and both the New York-based XRF and the lab returned results of “non-detect.” There was also one XRF measurement of a Kohl sample from Tunisia at a level of 712 ppm; the confirmatory lab testing of this sample is still pending.



There were four cosmetics samples in which the lab found extremely high lead concentrations ranging from 33,000 to 320,000 ppm (not included in the graph above), including a lipstick sample from the Philippines, a Kohl sample from Pakistan, yellow nail polish from Colombia, and an eyeliner (aka “chilo”) from Ghana. The XRF overestimated the lead concentrations in the two samples with the highest lead concentration by about a factor of 3. In the case of the yellow nail polish, the XRF underestimated the lead concentration by a factor of 10. These deviations could be the result of XRF user error, but suggested to the Quality Control Team that the XRF’s accuracy may simply diminish at extreme lead concentrations.

Staple Dry Foods

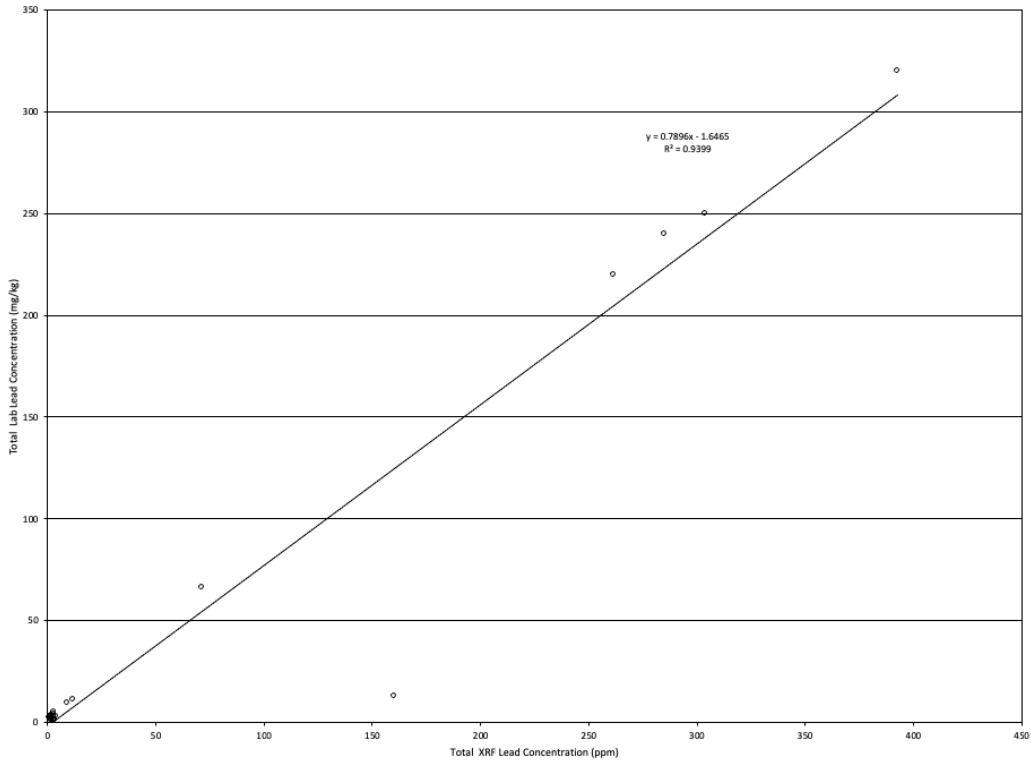
A total of 65 dry food samples representing local dietary staples were submitted for laboratory testing, including various grains, grain flours, and legumes. Laboratory results for 7 samples are pending. Lead was not detected in 48 of the samples reported, with 14 samples reported with lead levels below 1 ppm. There was only one sample with lead detected by the XRF at a concentration of 18 ppm with a corresponding laboratory report of 14 ppm for that sample. There was one sample for which the XRF measured 16.46 ppm for which the lab reported ND<0.1 ppm. Lastly, there were two samples for which the XRF did not detect lead (ND<2 ppm) for which the laboratory reported lead at 1.9 and 5.6 ppm.

Spices

A total of 125 spices were submitted for laboratory testing. Laboratory results for 11 samples are pending. The high number of spice samples sent to the lab reflects the fact that the reference value for spices is near the lower detection limit of XRFs. Types of spices sent to the lab included turmeric, chilis, tamarind, peppers, sunelli, and ginger, among many others. Lead was not detected by the laboratory in 44 of the samples. Lead was detected by the laboratory at concentrations of <2 ppm in 59 of the 125 samples and at concentrations of between 2 ppm and a high of 320 ppm in 18 of the 125 samples. For samples with laboratory lead concentrations >1 ppm, there was an excellent correlation between XRF and lab results, with a correlation coefficient (R^2) of 0.9399 including the results from Pakistan (see below), with the XRF results coming in slightly higher than lab results. Note that three samples with the highest lab results, ranging from 66 to 320 mg/kg, were turmeric samples obtained from 3 different markets in Tajikistan.

Note that lead was detected by XRF in 6 samples from Pakistan at levels ranging from 1 to 7 ppm in five of the samples and 160 ppm in the sixth sample. All six samples were submitted for testing at the Georgian Multitest laboratory. The laboratory reported lead from below the detection limit of <0.1 ppm to 0.49 ppm in the five samples with lower XRF results and 13.04 ppm in the sixth sample with an XRF reading of 160 ppm. Unfortunately, neither the XRF measurement nor the laboratory testing results could be independently verified at this time. As such, there is some uncertainty in the actual levels of lead in spice samples from Pakistan.

XRF vs. Laboratory Comparison
Total Lead in Spices

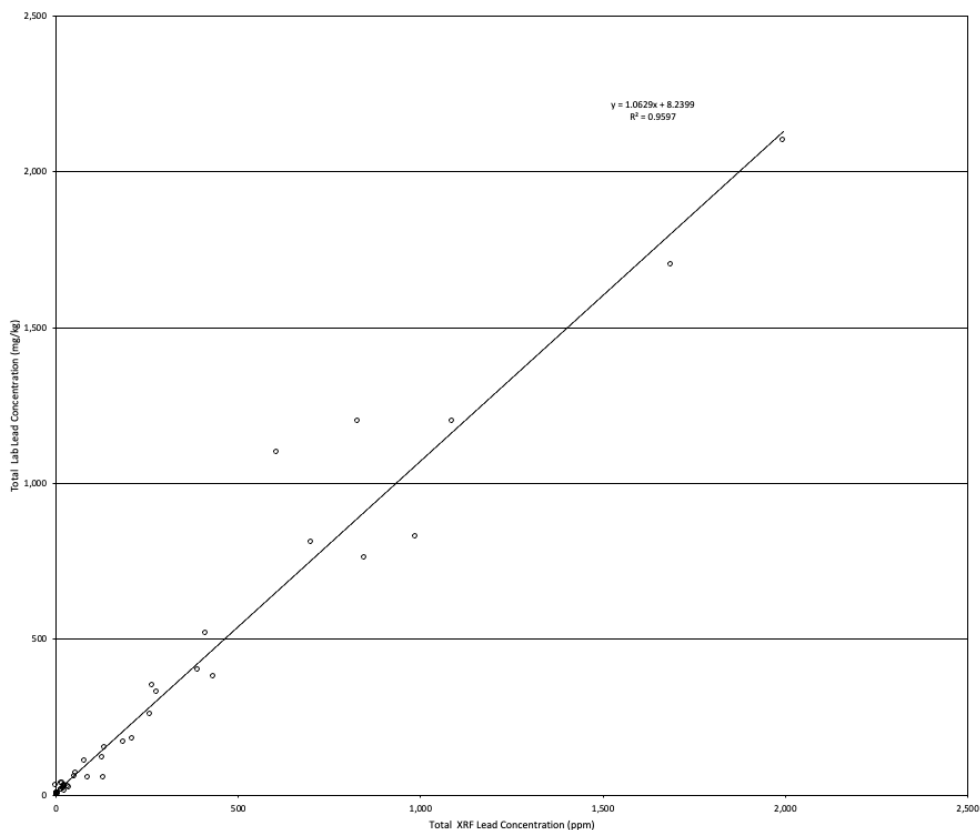


Toys

A total of 60 toys were submitted for confirmatory lab testing, including cars, stacking puzzles, dolls, and teething toys. Three of the toys were painted wood or metal, and the other 57 were plastic, typically with some surface color coating or colored plastic.

Lead was detected by the laboratory at concentrations of 0.77 to 2,100 ppm in 46 of the 57 plastic toy samples, with 25 plastic toys exceeding the 100 ppm threshold. There was excellent correlation between XRF and lab determined lead concentrations, with a correlation coefficient of 0.9597 for the plastic toy samples. Among the 11 toy samples for which the XRF did not detect lead, the lab analysis also reported no detectable lead (generally $ND < 0.50$ ppm).

XRF vs. Laboratory Comparison
Total Lead in Plastic Toys



There was one subcategory of toys for which XRF and laboratory results were not well correlated. Three painted toys were sent to New York for confirmatory lab testing (two wooden and one metal), each with relatively thin layers of paint. The paint was scraped off and submitted for laboratory testing for total lead for two of the toys. XRF measurements had indicated lead levels of 145 and 416 ppm, whereas the laboratory results indicated lead concentrations of 4,500 and 7,600 ppm, respectively, which is 20 to 30 times higher than the XRF readings. It appears that due to the very thin layer of paint, the XRF analyzed not only the paint, but likely also some part of the material under the paint. For example, the image below shows a thin layer of paint scraped from a wooden toy. The XRF value likely represents a combination of lead concentrations in the paint and the underlying wood (which likely contains no lead), resulting in a far lower value than the scraped paint alone. Conversely, one plastic elephant toy from India indicated lead at a level of 300 ppm using the XRF. The paint of the elephant was scraped off, and the plastic itself had an XRF lead concentration of 400 ppm, whereas lead was not detected in the scraped paint itself, also using the XRF. The plastic elephant or the scraped paint was not submitted for laboratory testing.



Sample 07SUR086 from Indonesia demonstrating thin red paint scraped from a wooden toy for laboratory testing.

Most toy samples were heterogeneous in nature, composed of differing parts, materials, and/or colors (e.g., a doll with a soft head on a hard plastic body), as shown below. Field XRF measurements of these heterogeneous toys did not necessarily represent all portions of these toys. The reported results represent maximum lead concentrations found for each toy.

Many of the toys shipped to New York for confirmatory testing contained internal metal parts (e.g., batteries, wires, gear boxes, lights etc.), as shown in the image below. XRF measurements taken in New York of selected toys suggested that the presence of these internal metal components could increase XRF readings taken on the surface of the toys. Such XRF readings of toys with internal metal parts were typically lower when the toys were disassembled and additional XRF measurements of individual external portions were tested without the metal components attached.



Heterogeneous toy samples from Kyrgyzstan.

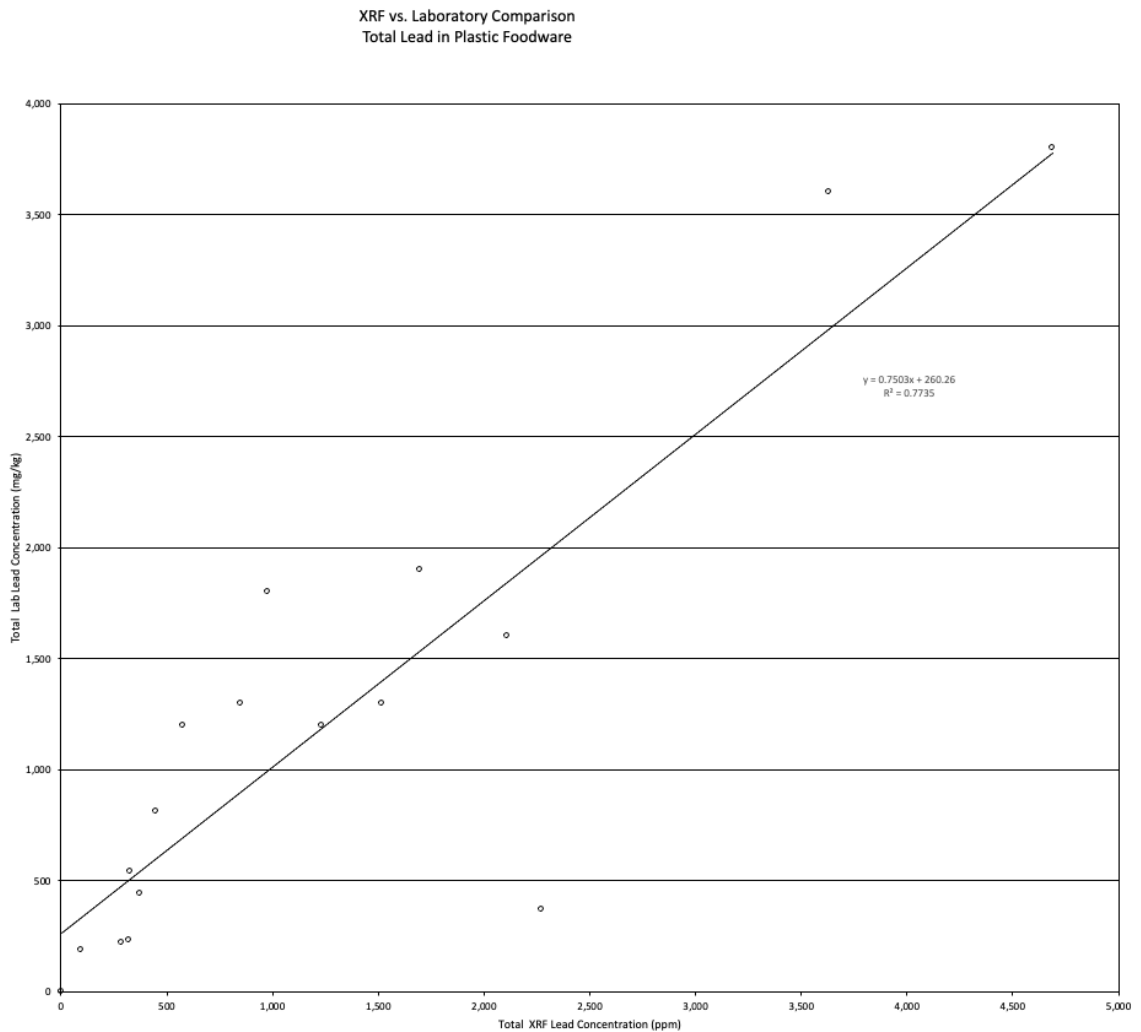


Sample 28BAK004 from Azerbaijan and 24ABU093 from Nigeria showing metal interior to toy.

Plastic Foodware

A total of 17 plastic foodware items were submitted for lab testing, including cups, bowls, plates, and a sweets container. Lead was detected in all 17 samples that were reported at levels between 1.4 and 3,600 ppm. There was a good correlation between XRF and lab results, with a correlation coefficient (R^2) of 0.7735.

One of the items tested was a small plastic container for sweets from Colombia, containing a type of a sweet called manjar blanco. Laboratory testing of the container itself indicated 1,300 ppm of lead. Lab testing of the sweets contained therein did not detect lead.



Other Foods

The Other Foods category contains predominantly sweets and some snacks (e.g., chips, pretzels, roasted crickets, pinole). A total of 15 samples were submitted for laboratory testing for total lead. Among these, two samples showed significant discrepancies between the XRF and lab results, and one of these discrepancies is likely explained by sample heterogeneity.

For 15 samples, neither the XRF nor the lab found detectable lead levels. In one sample, the laboratory reported more than 2 times the XRF measurement (18 ppm lab lead vs. 7 ppm XRF lead). This may have been due to heterogeneity of the sample, which was scraped from a ceramic container with 67,200 ppm of lead in the glazed portion and may have affected one part of the sweet more than others (see Ceramics section). In another sample, XRF readings from both the field and the New York-based XRF showed values between 6 and 10 ppm, but the lab did not detect lead.

Ceramics

A total of 93 ceramic samples were shipped to New York for additional quality control testing with the New York-based XRF, which generally confirmed the elevated levels reported by the country teams, with the differences related to the heterogeneity inherent in the ceramic items. Such heterogeneity included differences in the base glaze versus decorative patterns on various portions of the ceramics, as well as heterogeneity in the decorative patterns themselves.

Elevated lead levels in ceramics are often the result of lead-based glazes. The Quality Control Team determined that confirmatory testing of glazed ceramics by a lab was not a valuable exercise as the glaze itself could not be physically separated from the clay for an appropriate comparison. As an example of this challenge, a shard from a ceramic cup from Armenia with an XRF reading of 5,714 ppm was submitted for laboratory testing. The laboratory ground the shard into dust and analyzed the total lead concentration, indicating 260 ppm. The Quality Control Team concluded that the lab result was effectively diluted by the presence of unglazed interior clay in the sample and thus not representative of the surface levels that would influence exposure.

A number of ceramics samples were shipped to New York for possible leachate testing, which was not part of the official RMS protocol, but is an activity that Pure Earth is undertaking in conjunction with the RMS. One ceramic sample, a small glazed container holding a tamarind sweet, had an XRF lead concentration of 67,200 ppm in the glazed portion of the container. The tamarind sweet was scraped from the inside of the glazed container and submitted for laboratory testing. Lead was detected at a concentration of 18 ppm in the sample, indicating that the lead from the glaze is likely leaching into the tamarind sweet. Further leachate testing of other ceramics is warranted.

The following observations were made that are pertinent to the interpretation of XRF data for ceramic samples. The glaze in many samples was found to contain elevated lead concentrations. However, many ceramic samples sent to New York were heterogeneous in ways that could impact actual exposure levels. For example, a number of ceramic items were glazed white, with decorative patterns covering parts of the items, some interior and some exterior. In some cases, highly elevated lead concentrations were measured in these decorative patterns, but not in the white glaze which constituted the majority of the surface. The proportion and location of these decorative portions would impact an exposure assessment (e.g., interior vs. exterior; rim vs. body; small vs. large percentage). In addition, the leaching characteristics of decorative portions may be different than the white glaze. An observer may suggest that only interior portions of ceramics should be analyzed, but this decision would also invite complications. For one, some decorative portions spanned the interior and exterior. In other cases, the decorative portion was limited to the exterior, but included areas where a person's mouth would contact the cup. As an example, the cup from Peru shown in the image below was primarily covered in a white glaze that showed an XRF reading of ND<17 in this portion of the cup, while lead was detected by XRF at a level of 12,000 ppm in the decorative exterior portion of the cup, including where one's mouth would come into contact. There were a number of other such instances, including a ceramic mug from Nigeria with over 40,000 ppm of lead in a silvery band along the interior and exterior rim of the mug, but about 700 ppm in the remaining glazed portions.



Sample 17PUC077 from Peru with XRF lead at ND,17 in the interior white portion and 12,000 ppm in the decorative exterior portion.

Metallic Foodware

A total of 168 metallic foodware samples were shipped to New York for additional testing. New York staff analyzed 96 of these samples with the New York-based XRF, which confirmed the elevated levels reported by the country teams, considering variability in the aluminum items themselves (e.g., interior versus exterior, welded seams, rivets). 93 of the 154 pots sent to New York were selected for further leachate testing. While leachate testing is not an official component of the RMS protocol, it is an activity that Pure Earth deems extremely valuable for understanding the potential contribution to lead poisoning from metallic foodware, and which Pure Earth will continue conducting and reporting on through a separate report to follow. Samples of the metal itself were not submitted for confirmatory laboratory testing, as this would have been destructive to the pots themselves which were being leachate tested. As such, there is no comparison between XRF and laboratory results for metallic foodware.

The following observations are pertinent to the interpretation of XRF data for metallic foodware samples. The RMS database contains maximum lead concentrations for metallic foodware samples, but does not necessarily indicate which portion(s) were tested and the corresponding lead concentrations of each portion. A number of the metallic foodware samples appeared homogeneous, yet showed highly variable XRF readings in the side versus bottom or the inside versus the outside of the pot for example. Some pots contained paint or coatings on the exterior that had the highest lead levels using the XRF, while others had the highest lead levels in interior coatings. For example, an aluminum pan from Azerbaijan that had red paint on the exterior showed an exterior lead concentration of 8,457 ppm. However, lead was not detected in the interior coating of the pan, nor in the uncoated exterior bottom of the pan. This heterogeneity in metallic foodware has been observed by taking multiple XRF readings and will be helpful in interpreting the leachate testing that is summarized in the Metallic Foodware section.