

An Investigation of Lead in Traditional Eyeliners: An Important Source of Lead Exposure in Resettled Populations in King County, Washington

Aesha Mokashi

A thesis

submitted in partial fulfillment of the
requirements for the degree of

Master of Science

University of Washington

2024

Committee:

Diana M. Ceballos

Christopher D. Simpson

Katie M. Fellows

Program Authorized to Offer Degree:

Department of Environmental and Occupational Health Sciences

©Copyright 2024

Aesha Mokashi

University of Washington

Abstract

An Investigation of Lead in Traditional Eyeliners: An Important Source of Lead Exposure in Resettled Populations in King County, Washington

Aesha Mokashi

Chair of the Supervisory Committee:

Diana M. Ceballos

Department of Environmental and Occupational Health Sciences

Traditional eyeliners have been known to contain high concentrations of lead, a neurotoxic metal. In King County, Washington USA, several refugee, and immigrant populations continue to use these products on young children. Lead is toxic, even at low levels, especially in young children where it can cause neurological defects and growth delays. We investigated lead in eyeliners used by these communities, including traditional and nontraditional eyeliners, to determine eyeliner characteristics that resulted in higher lead concentrations. We also assessed and optimized Public Health- Seattle and King County's (PHSKC) current X-ray Fluorescence (XRF) analysis methods for cosmetic lead screening. We found that traditional eyeliners contained significantly greater median lead concentrations (10 ppm) than nontraditional ones (0.06 ppm), and eyeliners made in the United States and Europe had significantly lower median lead concentrations (0.94 ppm) than those made in Afghanistan (29 ppm) and other low- and middle-income countries (2.8 ppm). For eyeliner screening, we found that the current PHSKC XRF method is adequate for screening, although it overestimated lead concentrations in eyeliners. The XRF Soil method we developed may measure lead in eyeliners closer to the laboratory-derived values than the current method, but it requires further validation. Some "lead-free" labeled traditional eyeliners contained dangerously high lead concentrations (up to 610,000 ppm) well above the FDA guideline of 10 ppm and WA safety guideline of 1 ppm. Our findings suggest that traditional eyeliners being used in the United States are still produced with hazardous materials.

Introduction

Traditional eyeliners, known as kohl, kajal, or surma, have been used for thousands of years by cultures worldwide to beautify, treat, and protect the eye. Over the last few decades, researchers have found that many of these products are made with or contaminated by toxic ingredients including lead. Despite these findings, communities continue to apply these eyeliners, including on children in King County. As a result, we want to understand better the products our local community uses to develop guidance on safer practices.

From 2016-2020, the Washington Department of Health (DOH) observed that Afghan refugee children ages 0-16 had the highest prevalence of blood lead levels (BLLs) greater than 5 micrograms of lead per deciliter of whole blood ($\mu\text{g}/\text{dL}$) compared to all other refugee populations in Washington state (WA).⁽¹⁾ The Centers for Disease Control and Prevention (CDC) uses a blood lead reference value (BLRV) to identify children with BLLs that are higher than most children in the United States. The BLRV is based on the 97.5th percentile of BLLs in U.S. children ages 1-5. The current BLRV is 3.5 $\mu\text{g}/\text{dL}$; however, at the time of testing (2016-2020) the BLRV was 5 $\mu\text{g}/\text{dL}$.⁽²⁾ Per CDC guidelines, all incoming refugee children \leq 16 years of age should be tested for lead.⁽³⁾ During the four years, 44% of newly arrived Afghan children in Washington had BLLs greater than the previous BLRV of 5 $\mu\text{g}/\text{dL}$.⁽¹⁾

In response, Public Health- Seattle and King County's (PHSKC) Residential Services Program (RSP) in the Hazardous Waste Management Program (Haz Waste) conducted in-home environmental investigations in the residences of children with BLLs above the BLRV. They found that traditional eyeliners, often referred to as kajal, kohl, or surma, brought over from Afghanistan contained hazardous levels of lead.⁽⁴⁾ In addition, in 2022 and 2023, PHSKC held "Community Product Lead Testing Events" in collaboration with PHSKC's Environmental Health Community Toxics Section (EH CTS) and several community organizations, where they discovered multiple traditional eyeliners with high lead concentrations. In 2022, eight traditional eyeliners were determined to have lead concentrations ranging from 390,000 parts per million of lead (ppm) to 840,000 ppm.⁽⁵⁾ These eyeliners exceeded the United States Food and Drug Administration's (FDA) recommended maximum level of lead in cosmetics of 10 ppm, as well as recent legislation passed in WA that restricts lead or any lead compounds intentionally added or present, to 1 ppm or lower in any cosmetic.^(6,7) The discovery of these eyeliners suggests that multiple resettled populations in King County, WA regularly use cosmetics with high lead levels. No level of lead is safe for humans.⁽⁸⁾

Use of traditional eyeliner

Traditional eyeliners are commonly used throughout Africa, South Asia, and the Middle East. The use of traditional eyeliners dates back to the ancient Egyptians.⁽⁹⁾ Holy Islamic texts describe the Prophet Mohammed wearing eyeliner daily, and followers are encouraged to do so for religious practices and to increase eye health.^(10,11)

Today, traditional eyeliners, or surma as they are called in Afghanistan, are worn for various reasons, including aesthetic, religious, and medicinal purposes.⁽¹²⁻¹⁴⁾ Men, women, and children apply these products, but children are particularly vulnerable to lead exposure and lead absorption due to their propensity for hand-to-mouth behavior.⁽¹⁵⁾ While ingestion is one route of exposure, there is little information on the potential for lead to enter the bloodstream ocularly, via the mucous membranes in the eyelids.⁽¹⁶⁾

Afghan parents often apply surma to the children's outer and inner eyelids to emphasize the beauty of the eyes, ward off "the evil eye," or provide medicinal benefits. It is also important to acknowledge the role that familial and cultural values play in their application. For example, a study in Pakistan found that many new mothers applied eyeliner to their children because they had been counseled by older women in their families to do so.^(13,14) The passing down of eyeliner practices over generations, highlights how deeply rooted the use of traditional eyeliners is in tradition, religion, and culture.

Traditional eyeliner ingredients and terminology

Traditional eyeliner ingredients vary in part by their preparation (**Table 1**). Some traditional eyeliners are prepared by grinding antimony/stibnite (Sb_2S_3) to a fine powder; however, the rarity of this element has made many producers turn to galena (PbS), a lead-based mineral that is less expensive and easier to obtain.^(9,17) Alternatively, some communities in India and Pakistan burn plant oil using traditional candles (diya) and mix the resulting soot with oils such as clarified butter (ghee), or almond oil.^(18,19) These processes vary depending on the region. The elemental composition of traditional eyeliners depends on its texture. Hardy et al. (2002) and Filella et al. (2020) found that traditional powder eyeliners were often composed of a lead sulfide compound- most likely galena (**Table 10**).^(21,22) Cream or liquid traditional eyeliners tended to be composed of carbon compounds, silicone compounds, and waxes.^(21,22)

There is little regulation regarding traditional eyeliner manufacturing, and products can be readily sold and purchased in local shops or online. Some traditional eyeliners have labels that market the product as lead-free, despite analysis showing they contain high concentrations of heavy metals.⁽²⁰⁻²³⁾

The terms used to name traditional eyeliners vary based on the region and culture (**Table 1**). “Kohl” is often used to describe traditional eyeliners in the Middle East and other majority Islamic countries.⁽²⁴⁻²⁷⁾ Literature and online sources suggest that kohl is typically made by grinding minerals such as galena.^(17,24,28,29) The term “kajal” is predominantly used to describe traditional eyeliners used in India, and is often made with soot mixed with animal fats or vegetable oils.^(18,30) The term “surma” is used in Afghanistan, as well as North India and Pakistan.^(18,31) Some sources suggest that surma is made by grinding minerals,^(27,32,33) while others suggest it is made from soot.^(18,19) In Afghanistan, there are two types of surma used. One is “oily surma” which is made from combining soot and animal fat, the other is “stone surma” which is made by grinding minerals.⁽¹²⁾ The type of surma is an important distinction to make as it could impact the amount of lead that an Afghan child is exposed to. Multiple research articles consider kohl, kajal, and surma to be similar products, and these terms are used interchangeably.^(25,27,34) Because the terminology is inconsistent, a consumer cannot determine whether a product represents a lead exposure hazard based on its name alone.

Table 1: Description of traditional eyeliners based on region of use.

Term	Origin/Translation	Region of use	Common ingredients	Other names
Surma	Urdu word for antimony ⁽³⁵⁾	Afghanistan ^(12,33)	Naturally occurring ore ⁽³³⁾ ; Soot from burning animal fat or almonds ⁽¹²⁾	Stone surma (Afghanistan)- if product is made from grinding minerals ⁽¹²⁾
	Derived from Persian word “sormeh” ⁽²⁴⁾	Northern India ⁽¹⁸⁾ Pakistan ⁽³¹⁾	Grinding lead sulfide (PbS) also known as galena ⁽³²⁾	Oily surma (Afghanistan)- if product is made from soot and animal fat ⁽¹²⁾
Kajal	Arabic origin ⁽³⁶⁾	India ⁽³⁰⁾ Pakistan ⁽³¹⁾	Soot of sandalwood or other plants, mixed with oil or clarified butter (ghee) ^(18,30,37)	Kahal, Kanmashi, Kannu, Kappu, Kaatuka, Kanmai ⁽¹⁰⁾
Kohl	Arabic origin ⁽²⁸⁾	Middle East ^(28,32,37)	Primarily galena - can be mixed with herbs (saffron, neem, fennel, etc.), gemstones, and/or marine life (coral, pearl, etc.) ^(28,30,37)	Kohl stone ⁽²⁸⁾

Lead in cosmetics- adverse health effects and regulation

In children, lead exposure is particularly of concern when it comes to neurological development.⁽¹⁵⁾ Adverse neurological health effects include changes in mood, decreased learning ability, deficiencies in neuromotor and neurosensory functions, and encephalopathy.⁽¹⁵⁾ Lead may also interfere with development in the womb leading to decreased birth weight and size in infants.⁽¹⁵⁾ A few epidemiological studies found a significant association between the application of traditional eyeliners

in young children and having BLLs above the BLRV; however, causation was not established through more rigorous testing like determining the specific lead isotope in the blood.^(16,38-40) Two relatively recent clinical cases observed children with BLLs above the BLRV due to the application of traditional eyeliner, followed by a decrease in BLLs once application ceased.^(41,42)

A scientific statement recently put out by the American Heart Association (AHA) explains that low-level chronic exposure to lead can also result in ischemic heart disease, heart rate variability, stroke, and peripheral artery disease.⁽⁴³⁾

Resettled children are especially prone to adverse health effects due to their forced displacement, experiences with conflict, and statelessness.⁽⁴⁴⁾ A review of BLLs of refugee children found that they had a higher prevalence of cases that exceeded the BLRV compared to the general population.⁽⁴⁵⁾ Resettled populations often arrive from low- and middle-income countries (LMICs), and studies have shown that children in LMICs with lead poisoning have a greater loss of intelligence quotient (IQ) and greater incidence of mild mental retardation (MMR) compared to children in higher income countries.^(46,47) Therefore, Afghan children in the United States are especially vulnerable to the effects of lead.

Despite the in-depth documentation on the hazards of lead, regulation on lead in cosmetics varies from country to country (**Table 2**).

Table 2: Guidance on lead limits in cosmetics within the United States and globally.

Country	Lead limit (ppm) ^a	Specific to eyeliners?	Reference/Regulation	Comments
<i>United States</i>				
Federal	10	No- refers to lip products and any externally applied cosmetics	Lead in Cosmetic Lip Products and Externally Applied Cosmetics: Recommended Maximum Level ⁽⁶⁾	10 ppm is a guidance for recommended maximum level of lead in cosmetic products by the FDA ^b FDA Import Alert 53-06: The FDA has issued an Import Alert on “Kohl, Kajal, Al-Kahal, Surma, Tiro, Tozali, or Kwalli” ⁽²⁰⁾
Washington state (WA)	1	No- refers to any cosmetics as defined in RCW 69.04.011 ⁽⁴⁸⁾	Washington HB 1047 ⁽⁴⁹⁾	This bill restricts lead or any lead compounds intentionally added or present at 1 ppm or higher in any cosmetic (goes into effect in 2025)
<i>Other countries/regions</i>				
European Union	0	No- refers to all cosmetics.	(EC) No 1223/2009 (Article 14, Annex 2) ⁽⁵⁰⁾	Non-intended presence of a small quantity of a prohibited substance (such as lead and its compounds) is permitted provided that it follows certain requirements ⁽⁵⁰⁾
Germany	5	Yes- refers to kajal, eye shadow, eyeliner, makeup powder, rouge, as well as theater, fan, or carnival makeup	Technically avoidable heavy metal contents in cosmetic products ⁽⁵¹⁾	Heavy metal content in traditional eyeliners exceeding 5 ppm is considered technically avoidable. For all other cosmetics, the limit is 2 ppm.
Canada	10	No- refers to any cosmetic	Guidance on Heavy Metal Impurities: Section 4 ⁽⁵²⁾	Pb content above 10 ppm is considered technically avoidable
India	20	No- applies to any ingredient meant to add color to a cosmetic	Rule 134 of the Drugs and Cosmetics Act and Rules ⁽⁵³⁾	“The permitted synthetic organic colors and natural organic colors used in the cosmetic shall not contain more than 20 ppm of lead” ⁽⁵³⁾
Saudi Arabia	10	No- refers to all cosmetics and personal care products	NO. SFDA.CO/GSO 1943:2021 ⁽⁵⁴⁾	Non-intended and technically unavoidable levels of lead are permitted in cosmetics and personal care items if they do not exceed 10 ppm
Pakistan	No regulations found	N/A	N/A	
Afghanistan	No regulations found	N/A	N/A	

a. parts per million

b. United States Food and Drug Administration

In the United States, the FDA has placed an Import Alert on products labeled “kohl, kajal, al-kahal, surma, tiro, tozali, or kwalli” and has identified these products as “illegal color additives.”⁽²⁰⁾ For all cosmetics, the United States FDA has placed a recommended maximum lead concentration of 10 ppm.⁽⁶⁾ WA passed legislation in 2023 that limits the concentration of lead or its compounds to 1 ppm or lower.⁽⁴⁹⁾ The European Union (EU) has stated that lead and its compounds are prohibited at any concentration in all cosmetics; however, the non-intended presence of small quantities is allowed provided that the lead concentration follows certain requirements.⁽⁵⁰⁾ In Germany, lead in traditional eyeliners cannot exceed 5 ppm.⁽⁵¹⁾ Germany is the only country with lead regulations that are specific to traditional eyeliners. In India, where many traditional eyeliners are manufactured and used, any permitted product used to color a cosmetic must not contain more than 20 ppm lead.⁽⁵³⁾ In Saudi Arabia, practicing Muslims often purchase traditional eyeliners here because they add holy water known as “zam zam”, which can only be found in Mecca.⁽⁵⁵⁾ Saudi Arabia has set a lead limit of 10 ppm in cosmetics and personal care items.⁽⁵⁴⁾ Many traditional eyeliners are also manufactured and worn in Pakistan, but the country does not have any regulations on lead in cosmetics. The same is true for Afghanistan. Despite being banned in the U.S., in-home visits by Haz Waste revealed that communities continue to use these eyeliners.⁽⁴⁾ The King County/Seattle Afghan community has been informed of the high levels of lead in their cosmetics, and as a result, they have begun using alternatives that they believe are lead-free. However, the safety of these products has not been assessed.

Previous Studies

Multiple researchers have measured lead in traditional eyeliners (**Table 9**). McMichael and Stoff (2017) analyzed ten samples of Afghan-produced traditional eyeliner using inductively coupled plasma mass spectrometry (ICP-MS).⁽³³⁾ The researchers determined that all samples contained lead with concentrations ranging from 400-830,000 ppm. Studies conducted in the Middle East found lead in traditional eyeliners ranging from 0-973,000 ppm.^(22,23,26,27,56) Researchers in Europe found similar results, with some traditional eyeliners containing little to no lead, and others containing thousands of ppm of lead.^(17,21) New York City’s Department of Health and Mental Hygiene has tested various cosmetics since 2011 including traditional eyeliners during in-home investigations and store surveys.⁽⁵⁷⁾ The results of these analyses have been published in an open database for anyone to access. PHSKC has developed a similar public database where all products tested at Community Product Lead Testing Events are cataloged.⁽⁵⁸⁾ New York City and King County have tested traditional eyeliners that contained lead in the percentage levels (>100,000 ppm).

Several researchers have also studied lead in nontraditional eye cosmetics (**Table 10**). Nontraditional eye cosmetics in this project refer to eye cosmetics that are not labeled as kajal, kohl, or surma. They are not made from ingredients commonly associated with traditional eyeliners such as galena, soot, and/or animal and vegetable oils. Nontraditional eye cosmetics include eyeliners, eyeshadows, and mascaras. Overall, nontraditional eye cosmetics had lead concentrations far below traditional eyeliners; however, many still contained lead levels that violated cosmetic regulations. In the United States, King County, WA, and New York City have found products that fall above the federal regulation of 20 ppm.^(57,58) Charter et al. (2011) in Canada found samples that had 110 ppm despite Canada's limit of 10 ppm lead in all cosmetics.⁽⁵⁹⁾ Three European studies found lead concentrations in eyeshadows fell below 20 ppm; however, the EU has prohibited lead contamination in all cosmetics, and Germany has stated that it is technically feasible to have lead concentrations <2 ppm.⁽⁶⁰⁻⁶²⁾ In Saudi Arabia, where the lead in cosmetics cannot exceed 10 ppm, Al-Saleh et al. (2009) found cosmetics that had lead contamination ranging from 0.52 to 58.72 ppm lead.⁽⁶³⁾ These studies show that despite having lead levels lower than traditional eyeliners, nontraditional cosmetics still contain hazardous lead levels, and consumers must think carefully about the products they use.

An important aspect to look at in these studies is the level of community engagement (**Table 9**) (**Table 10**). When assessing cosmetics for impurities, the study's relevance increases if the researchers test cosmetics used by community members. A few researchers looked at cosmetics that they established were used by the community. Charter et al. (2011) asked six women of varying ages to submit cosmetics they used regularly to be tested for heavy metals.⁽⁵⁹⁾ PHSKC-produced data came from samples tested at Community Product Lead Testing Events and at in-home investigations.⁽⁵⁸⁾ Both events have close collaboration with the local community. Similarly, data produced by New York City's Department of Health & Mental Hygiene from lead poisoning investigations came from products that families used regularly.⁽⁵⁷⁾ Hardy et al. (2002) looked at one eyeliner brought in by a concerned student, but the research team purchased the rest.⁽²²⁾ In general, most samples analyzed in previous literature were purchased by researchers. For studies conducted on traditional eyeliners in countries where they are commonly used, it is possible that the researchers purchased products they used regularly, but this has not been mentioned in any of the methods.

Several techniques are commonly used for analyzing samples, including inductively coupled plasma techniques (ICP), atomic absorption spectroscopy (AAS), and various X-ray techniques including X-ray fluorescence (XRF) analysis. The analysis technique is important to understand because accuracy and

precision can vary depending on the technique. ICP and AAS techniques are often referred to as the “gold standard” as they are theoretically able to achieve detection limits for lead in the part per billion (ppb) range.⁽⁵⁸⁾ The precision and accuracy ICP and AAS can achieve are important when comparing lead concentrations in products to regulatory limits. A combination of two or more EPA-developed methods must be used when determine whether an eyeliner falls within regulation to prevent false negative or false positive readings.⁽⁶⁴⁾

XRF analysis, on the other hand, is a non-destructive instrument often used in the field, as it is portable, and can rapidly quantify the elemental composition of a whole sample.^(21,58) XRF measurements cannot be used for regulatory purposes because the handheld instrument has relatively high detection limits.⁽⁵⁸⁾ Because XRF analyzers act as screening tools, they must quantify lead accurately, especially at concentrations close to regulatory limits. Traditional eyeliners are unique because they have a variety of textures, unique ingredients, and varying ranges of lead. Therefore, unless traditional eyeliners are tested with the optimal settings, the XRF may be unable to measure lead accurately, even at higher concentrations.

This Study

The aims of this study are as follows. In Aim 1 we want to know what eyeliners our local community members use and understand what factors result in greater lead concentrations in eyeliners. In Aim 2 we will compare PHSKC’s current XRF method with the gold-standard laboratory measurements to assess the XRF’s performance. Once we understand how well PHSKC’s XRF performs, we want to see how we can optimize the settings so that XRF measurements are as accurate as possible; therefore, in Aim 3 we will evaluate three XRF methods to optimize the XRF’s ability to measure lead in eyeliners.

Methods

Sample collection

We collected traditional and nontraditional eyeliners for this project. For each product collected, we gathered information on the type, country of production, brand and manufacturer, texture, disclosed ingredients, and whether the product had “lead-free” labeling from the product’s packaging, the website the eyeliner was purchased on, and from community members who were donating their product. Here we outline the various methods of sample collection.

Eyeliners provided by community

Community Product Lead Testing Events

Community members brought personal items to events to be analyzed by staff for lead content either on-site by XRF analysis, or by a commercial laboratory. Staff collected information on the product from participants and any labels that were on the product. Information collected included product description, brand name, manufacturer, the country the product was made in, purchase location, and batch number if applicable. Of all the products tested at these events, only eye cosmetics were used for this study. Eye cosmetics included eyeliners, eyeshadows, and mascaras, as well as traditional eyeliners such as kajal, kohl, or surma.

Donations from non-profit organization

Samples were also collected through collaboration with the Afghan Health Initiative (AHI), a grass-roots non-profit organization in WA founded by current and former immigrants and refugees from Afghanistan.⁽⁶⁵⁾ Due to the high prevalence of Afghan children with BLLS above the BLRV, Haz Waste collaborated with AHI to research potential lead exposures and effective interventions.^(1,4) AHI visited the homes of 35 families to administer questionnaires, and to collect traditional eyeliners for testing. Consequently, 22 products were given to the Research Services Team in the Haz Waste Program (Research team) for testing. AHI staff collected information from participants on the name of the eyeliner, where it was purchased, and whether it was homemade or not.

In-home investigations

PHSKC investigators found two traditional eyeliners that contained hazardous lead concentrations during in-home visits. These samples were then transferred to the Research team for further analysis.

Other community-recommended products

Two additional products were recommended for testing by other community members. One was recommended by the Multnomah County Health Department in Portland, Oregon, and another was recommended by a PHSKC employee. Kohl samples from Saudi Arabia were brought back for testing by a PHSKC employee. Since some products used by the Afghan community are purchased or made in Saudi Arabia, the samples donated by the PHSKC employee were relevant to this study.

Eyeliners purchased by researchers

Amazon

Many Afghans were aware of the dangers of lead in traditional eyeliners and were using cosmetics they deemed to be “safer alternatives” instead. They often purchased these products from local King County grocery stores or via online retailers such as Amazon. AHI provided the Research team with a list of these products, which were purchased from Amazon to test whether they were safer alternatives. The list included eyeshadow pallets, eyeliners, and mascaras. Some products were not found online, and as a result, we attempted to find alternatives that most resembled them.

Etsy and eBay

To assess the availability of hazardous products from online retailers other than Amazon, products were also purchased on eBay and Etsy. Search terms included “Surma,” “Kajal,” and “Kohl.” The products identified in the top search results were purchased as these were the most popular items on the site.

Traditional eyeliners purchased online often used two or even the three names of traditional eyeliners in their advertising; however, when categorizing traditional eyeliners, we only used the term displayed on the physical packaging.

Sample preparation and analysis

Samples were first prepared by the Research team for XRF analysis before being sent to a laboratory for chemical analysis.

Sample preparation

Products analyzed by our Research team were delivered to the Haz Waste program’s laboratory in Seattle, WA. They were stored at room temperature (68 C).

A 210-millimeter (mm) spatula (LevGo 17211 smartSpatula Disposable Polypropylene Laboratory Spatula, 210 mm (Pack of 300), Blue), was used to transfer eyeliner product from its original container into a plastic sample bag (Plymor Zipper Reclosable Plastic Bags, 2 Mil, 2" x 3" (Pack of 100): Industrial & Scientific). Sample bags were purchased so that their thickness did not interfere with the XRF's attenuation as suggested by Bruker. ⁽⁶⁶⁾ Bruker also recommended that samples should be at least ¾ inches thick when analyzed because the X-rays only penetrate up to ¾ inches when measuring soil-like samples.⁽⁶⁶⁾ While the goal was to analyze at least 1 g of product that was ¾ inches thick, six eyeliner pencils contained less than 1 gram of sample and may not have had enough sample to reach ¾ inches. We did not weigh the samples in the lab. We looked at the weight provided on the website and purchased enough products to sum to 1 gram. See **Appendix A** for detailed sample preparation.

Sample analysis

Once samples were collected and prepared, they were tested using XRF, laboratory chemical analysis, or both. Here we outline all methods used to analyze products for lead (**Table 3**). Samples were submitted for confirmatory chemical analysis to two laboratories: the University of Washington's (UW) Environmental Health Laboratory and Trace Organics Analysis Center (EHL TOAC, Seattle, WA), and a commercial laboratory (NVL Laboratories, Seattle, WA). Samples were measured using multiple laboratory techniques (ICP-MS, ICP-OES, GFAA) because we used pre-existing data collected by PHSKC during in-home investigations and Community Product Lead Testing Events who were not worried about testing consistency.

Table 3: Experimental design table showing the various methods of analysis used to test traditional and nontraditional eyeliners for lead.

Aim	Number of eyeliners	Number of samples analyzed per analysis method ¹	Method of analysis
Eyeliner characteristics (Aim 1)	143	11	ICP-MS
		68	ICP-OES
		64	GFAA
Comparing current eyeliner testing methods (Aim 2)	37	9	ICP-MS
		28	GFAA
		37	XRF RM Standard
XRF method optimization (Aim 3)	37	4	ICP-MS
		8	GFAA
		37	XRF RM Standard
		12	XRF RM 180
		12	XRF Soil

¹ This number may add up to more eyeliners than were tested because most eyeliners were analyzed using multiple methods.

XRF testing

To understand how well the standard method for XRF analysis compared to the gold standard laboratory methods, we analyzed a group of eyeliners and a group of calibrants using various XRF methods (**Table 3**).

Calibrant selection

Calibrants were selected on their similarity to traditional eyeliners based on previous literature and our initial findings via XRF (**Table 4**). The calibrants had to be powders, contain other elements typically found in traditional eyeliners, and they should have little to no arsenic - to prevent arsenic from overlapping lead spectra. We selected soil calibrants that covered the wide range of lead concentrations found in eyeliners, and the regulatory range for lead in cosmetics (i.e., 0-20 ppm) because the soil had similar texture and elements as many of the eyeliners we tested. Eleven calibrants were selected. Calibrants were purchased from Brammer Standard Company, Inc based in Houston, Texas.

Table 4: Soil calibrants used to assess the three XRF methods.

Calibrant name	Type/Class/Unit	Pb (ppm)	As (ppm)	Elements present that were also found in eyeliners	Other elements
GBM912-4	Multimetal ore/CRM/Powder (10 g)	5	2	Ag, Zn, Co, Cu, Ni	N/A
GBM322-11	Multimetal ore/CRM/Powder (10g)	15	0	Ag, Zn, Cu, Ni, S	N/A
GBM320-12	Multimetal ore/CRM/Powder (10g)	50	0	Ag, Zn Cu, Ni, S	N/A
GBM322-13	Copper ore/CRM/Powder (10g)	98	0	Ag, Zn, Cu, Ni, S	N/A
GBM904-3	Multimetal ore/CRM/Powder (10g)	489	0	Ag, Zn, Co, Cu, Ni	N/A
GBM999-8	Multimetal ore/CRM/Powder (10g)	1,061	185	Ag, Zn, Co, Cu, Ni, S	N/A
GBM922-15	Multimetal ore/CRM/Powder (10g)	4,723	0	Ag, Zn, Cu, Ni, S	N/A
GBM323-5	Multimetal ore/CRM/Powder (10g)	10,556	1257	Ag, Zn, Co, Cu, Ni	N/A
NCS DC73510a	Multimetal ore/CRM/Powder (50g)	51,300	1244	Ag, Al, Fe, Zn, SiO ₂ , Cu, Mn, Sn, Cd, C, S	Re, CaO, Ce, Ga, Ge, Hg, In, K ₂ O, MgO, Mo, Na ₂ O, Re, Sb, Se, Te, Tl, W
CAN CPB-3	Lead ore/CRM/Powder (100g)	580,000	391	Ag, Al, Co, Cr, Fe, Zn, Ca, Cd, Ni, Cu, S, SiO ₂ , Sn	Au, Ba, C, Hg, K, La, Mg, Na, Sb, Sr, Th, Y
NCS DC71309	Sulfur ore/CRM/powder (5g)	842,600	5.3	Ag, Fe, Zn, Cd, Co, Cu, S, Sn	Bi, Ga, Ge, In, Sb, Te, Tl

Testing with the XRF

The Bruker S1 Titan Handheld XRF Analyzer was used for analysis (Bruker Corporation, Billerica, Massachusetts). Before product testing began, five calibration checks were performed with the appropriate certified standard (SAC305 or CS-M5) to ensure that average lead concentrations fell within the accepted limits provided in the device's Certificate of Calibration (**Table 5**). Once this quality check was conducted, sample testing proceeded. Five replicate measures were taken in the same location for each sample. Once all samples had been measured, five calibration checks were conducted on the standards.

Restricted Materials Standard XRF method (XRF RM Standard)

Once samples were adequately prepared, sample testing was conducted with the XRF using the standard PKSKC method- XRF RM Standard. XRF RM Standard is used during in-home investigations, Community Product Lead Testing Events, and most research projects.

For XRF RM Standard, the 'Restricted Materials' setting was selected (**Table 5**). Restricted Materials detects restricted materials regulated under the 'Restriction of Hazardous Substances' directive. The device's 'Method' was set to 'Automatic Calibration Selection.' Automatic Calibration Selection utilizes Bruker technology to automatically select the correct analytical testing sub mode (e.g., metal, paint, plastic, etc.). The XRF measures elemental concentrations in two phases: Phase 1 measures heavier elements (such as lead), and Phase 2 measures lighter elements. Phase 1 was set to 30 seconds at 50 kilovolts (kV), and Phase 2 to 30 seconds at 15 kV.

Limits of detection (LOD) are important to understand when working with an XRF. LOD is the lowest concentration that the instrument can reliably identify the presence of an element. For XRF RM Standard, we determined the LOD using the error terms for soil calibrant measurements <LOD (**Table 4**). The minimum LOD for XRF RM Standard was 10 ppm lead and the average LOD was 11.1 ppm lead. The average LOD was calculated by taking the mean of all error terms for <LOD measurements. XRF RM Standard measurements were <LOD for all calibrants with ≤ 15 ppm lead.

Restricted Materials 180 s XRF Method (XRF RM 180)

The next method to test was Restricted Materials with a 180-second Phase 1 (XRF RM 180) (**Table 5**). XRF RM 180 was used based on Bruker's Restricted Materials Calibration sheet for the instrument. The LOD for the instrument was determined using 180-second phases because the instrument's LOD improves as

a function of the square root of the analysis time. Therefore, XRF RM 180 was configured to maximize the amount of time the XRF analyzes for heavy metals.

The LOD for XRF RM 180 was calculated by Bruker. The minimum LOD for XRF RM 180 was 4.5 ppm and the average LOD was 30.8 ppm (**Table 5**). XRF RM 180 measurements were <LOD for calibrants with 5 ppm lead.

XRF Soil Method

The final method was the XRF Soil method (**Table 5**). This method was developed to measure lead in soil accurately; we considered powdered eyeliners to be similar in texture and composition to soil. We calculated the minimum and average LOD for XRF Soil using the error terms for calibrant measurements <LOD. The minimum LOD and average LOD were 14 ppm (**Table 5**). XRF Soil was the only method that detected lead in calibrants; however, only one of the five measurements was above the LOD.

Table 5: The three methods used to optimize XRF analysis of lead in traditional eyeliners. We compared the standard method (XRF RM Standard) that PHSKC uses during in-home investigations, Community Product Lead Testing Events, and research projects, with two new methods- RM 180 and Soil. Method name should not be confused with the Device’s ‘Method.’

Method name	Minimum LOD (ppm) ¹	Average LOD (ppm) ²	Setting	Device’s method	Phase 1	Phase 2	Standard
XRF RM Standard	10 ¹	11.1 ¹	Restricted Materials	Automatic Calibration Selection	30 s, 50 kV	30 s, 15 kV	SAC305
RM 180	4.5 ²	30.8 ²	Restricted Materials	Automatic Calibration Selection	180 s, 50 kV	0 s	SAC305
Soil	14 ¹	14 ¹	Soil	Soil Contaminant-Nutrient	60 s, 50 kV	0 s	CS-M5

¹ LOD for XRF RM Standard and Soil was determined using the calibrant measurements. Minimum LOD was the lowest error term of all <LOD calibrant measurements. Average LOD was determined by taking the average of all error terms for <LOD measurements.

² Bruker determined the minimum and average LOD for RM 180.

Inductively Coupled Plasma Mass Spectrometry (ICP-MS)

The UW’s EHL TOAC conducted ICP-MS with an Agilent 7900 SPS4 autosampler. Samples were analyzed using the EHL SOP-07 method (based on the EPA 6020a Rev.1 2007 method).⁽⁶⁷⁾ Reagents and method blanks were run with each batch, and samples were blank-corrected afterward for quality check. Two sub-aliquots per sample were analyzed to demonstrate homogeneity/heterogeneity. Matrix spike recovery was performed on one sample. The relevant detection limit was reported as Reporting Limit (RL). RL varies based on the sample weight and complexity of the sample matrix, and it is equal to the

limit of quantitation (LOQ) (personal correspondence with Shar Samy). The RL for our samples was 0.35 ppm (**Appendix B**). The RL was determined by taking the lowest calibrant used on the instrument (35 nanograms (ng)) and dividing by a representative sample aliquot mass (100 milligrams (mg)) to get 0.35 ng/mg (0.35 ppm). 11 samples were analyzed using this method, which were a combination of items donated to and purchased by the Research team (**Appendix B**).

Inductively Coupled Plasma Optical Emission Spectroscopy (ICP-OES)

ICP-OES was conducted by NVL laboratories using the CPSC-CH-E1002-08 (Children's non-metals Products) method.⁽⁶⁸⁾ For quality control, samples were prepared with deionized water and high-purity acids to prevent the presence of background lead. Each sample set was analyzed with a blank to check for background levels (personal correspondence with Shalini Patel). Blank corrections were not necessary because none of the blanks contained lead levels above the RL. The ICP instrument detection limit was 0.01 ppm. The average RL for samples <RL was 14.75 ppm (**Appendix B**). We calculated an average RL because RL differed for each sample due to weight and matrix differences. RL at NVL is based on the lowest concentration analyzed daily which can be quantified with a certain degree of confidence and can be reproducible within an established acceptable range (personal correspondence with Nick Ly). Sixty-eight products collected at Community Product Lead Testing Events or purchased by the Research team were analyzed using ICP-OES. Fifty-three of the samples were <RL and 32 were non-detects (ND) (**Appendix B**).

Graphite Furnace Atomic Absorption Spectrometry (GFAA)

NVL Laboratories conducted all GFAA analyses. Samples were analyzed using the EPA 7010 method.⁽⁶⁹⁾ Samples were prepared with deionized water and high-purity acids to prevent the presence of background lead. Each sample set was analyzed with a blank to check for background levels (personal correspondence with Shalini Patel). Blank corrections were not necessary because none of the blanks contained lead levels above the RL. The GFAA instrument detection limit for this method is 0.01 ppm. The RL for 1 g of sample was 0.2 ppm, and the RL for 0.5 g of sample was 0.4 ppm. We averaged all samples that were <RL to obtain a mean RL of 0.3 ppm (**Appendix B**). GFAA was chosen because its RL was below 1 ppm. New WA regulation limits lead in cosmetics to 1 ppm or less, which was why it was important samples were analyzed using GFAA because its RL for lead is below 1 ppm (**Table 2**).⁽⁷⁾ 66 products collected at Community Product Lead Testing Events or purchased by the Research team were analyzed using GFAA (**Appendix B**). Eight products were <RL and two were ND (**Appendix B**).

Statistical analysis

Various statistical analyses were performed using Excel Version 2405 (Build 17628.20110) and RStudio Version 2024.04.0 Build 735. We set the significance to $p < 0.05$.

Left-censored imputation and maximum likelihood were used for censored observations.

206 499 0657

Eyeliner characteristics

We proposed to investigate whether certain eyeliner characteristics were associated with significantly higher lead concentrations. Bivariate and multivariate analyses were conducted using nonparametric tests to determine whether there were significant differences ($p < 0.05$) in lead concentrations. For bivariate analysis, the 2-sided Wilcoxon Rank Sum test was used. For multivariate analysis, the Kruskal-Wallis test was employed. This method was derived from Ceballos et al. 2021.⁽⁷⁰⁾ To ensure our analysis had enough power, we confirmed that each variable had a minimum of six products. Our data was left-censored because 32 samples were ND (**Appendix B**). Using NADA2 and EnvStats packages in R, left-censored imputation and maximum likelihood were used for censored observations to conduct descriptive analyses. This method was based on a method provided on the website rdrv.io and through personal correspondence with Cynthia Wu.⁽⁷¹⁾

Traditional and nontraditional eyeliners were compared. Traditional eyeliners were defined as any product that had the words “kajal,” “kohl,” or “surma,” on its packaging or had been specifically described as a traditional eyeliner by its user. Nontraditional eyeliners included mascaras, eyeshadows, and eyeliners brought to Community Product Lead Testing Events, and products reportedly used by Afghan community members. We tested eyeliners used by the Afghan community to determine potential safer alternatives that the community would be comfortable adopting.

Traditional eyeliners were further evaluated by comparing the three traditional eyeliner types: “kajal,” “surma,” and “kohl.” These products have been found to contain hazardous lead concentrations in the Afghan community and at Community Product Lead Testing Events.⁽⁵⁸⁾ Labels on the eyeliner’s packaging were used to determine the type of eyeliner. If no label was present on the packaging, descriptions provided by the owner were used to determine the eyeliner type.

Products made in the United States and Europe, Afghanistan, and “Other” countries (China, Egypt, Ethiopia, India, Morocco, Mexico, Pakistan, and Saudi Arabia) were evaluated. The US and Europe have similar strict regulations for cosmetics manufacturing, and therefore American and European-made products were grouped into, “US/Eu.” The country of production was determined either based on the eyeliner’s packaging or on the website it was purchased on. If this information was not available, the production country was considered “unknown”.

Powder and cream eyeliners were compared next. Powdered eyeliners are more likely to be made by grinding galena, and cream eyeliners have a greater likelihood of being made from soot (**Table 1**), therefore we wanted to statistically compare the two categories to assess whether this was the case with our products. The texture of the eyeliner was determined visually. Eyeliners were considered a powder if they were dry, and their particles easily separated from each other. Eyeliners were considered a cream if they were wet, and the product did not easily separate into particles. Liquids such as mascaras and liquid eyeliner were considered creams because they were wet rather than dry. Some products from Community Lead Product Testing Events did not provide texture details and were labeled as “unknown”.

Eyeliners with “lead-free” labels on their packaging were compared against eyeliners without. Previous studies found that many traditional eyeliners often had “lead-free” labels on their packaging despite containing hazardous lead concentrations.^(21,23) Eyeliners from Community Product Testing Events were considered “unknown” because they did not provide data on “lead-free” labeling.

Finally, homemade and manufactured eyeliners were compared. It was important to understand if certain practices conducted in the home or a potentially more regulated setting, might limit a product’s lead concentration. Homemade eyeliners were considered homemade if the website they were purchased on said they were homemade, or if the community member who donated them said they had made the product at home. Manufactured eyeliners were determined by examining eyeliner packaging or looking for a manufacturer on the eyeliner’s website. If the manufacturer was not stated, the eyeliner’s production method was considered “unknown”.

[Comparing current eyeliner testing methods](#)

We conducted linear regressions between PHSKC’s current XRF method (XRF RM Standard) derived lead concentrations, and the laboratory-derived lead concentrations to assess the ability of XRF RM Standard (Table 6).

Eyeliner measurements were <LOD were removed because the lead concentrations fell below the instrument’s ability to detect lead, and therefore it was unable to estimate a lead concentration in the eyeliner without high error. A separate analysis that included <LOD is included in the appendix (Appendix C). The average LOD for XRF RM Standard was 11.1 ppm (Table 5). Laboratory-derived lead concentrations that fell <RL were kept in the analysis because the instrument detected lead in the product within a certain degree of confidence.

The categories for statistical analysis were based on our observations of the XRF’s performance (Table 6). Above 400,000 ppm lead, the XRF sensor appeared to be over-saturated as the number of absorbing species increased, resulting in a curvilinear relationship; therefore, eyeliners below 400,000 ppm were investigated separately. No cream eyeliners were above 1,000 ppm; however, three outliers above 90 ppm lead were removed. Once descriptive analyses were conducted, Pearson’s and Spearman’s correlation coefficients were calculated, and linear regressions were performed. EPA SW-846 Method 6200: Determination of Elemental Concentrations in Soil and Sediment was used to assess XRF RM Standard’s performance.⁽⁷²⁾ XRF measurements for the certified standard were assessed for quality control. (Appendix E).

Table 6: The various categories of eyeliners that we compared XRF (XRF RM Standard) Pb measurements to laboratory (ICP-MS, ICP-OES, GFAA) Pb measurements to assess XRF performance.

Category	Sample number
All eyeliners	37
All eyeliners <400,000 ppm	23
Powder <400,000 ppm	10
Cream <90 ppm	8

XRF method optimization

We tested three XRF methods (XRF RM Standard, XRF RM 180, XRF Soil) and conducted linear regressions between XRF-derived measurements and the laboratory-derived lead concentrations to optimize XRF testing of eyeliners. XRF Soil was the only method that detected lead in the 5 ppm lead calibrant, therefore most XRF measurements for this calibrant were removed from our analyses. XRF RM Standard did not detect lead below 50 ppm lead. A separate analysis was conducted including <LOD measurements (Appendix D).

Descriptive analyses were conducted for each calibrant and XRF method. Once descriptive analyses were conducted, Pearson's and Spearman's correlation coefficients were calculated, and linear regressions were performed. For quality control, we assessed the XRF measurements for the certified standards (SAC305, CS-M5) and looked at the mean and standard deviations (**Appendix E**).

For the 12 eyeliners tested using all three XRF methods, none of the XRF methods detected lead below 32 ppm; therefore, <LOD measurements were removed. A separate analysis was conducted including <LOD measurements (**Appendix D**). The same statistical methods used for the calibrant testing were employed. None of the laboratory-derived lead concentrations were ND or <RL.

Results

Eyeliners characteristics

Nontraditional and traditional eyeliners

Traditional eyeliners had significantly greater concentrations of lead than nontraditional eyeliners ($p < 0.001$) (**Table 7**). Traditional eyeliners had a median of 10 ppm lead, and nontraditional eyeliners had a median of 0.06 ppm lead. 76.67% of nontraditional eyeliners were below the RL as compared to 20% of traditional eyeliners.

Traditional eyeliner names: kohl, kajal, surma

When comparing 26 kohl samples, 27 kajal samples, and 20 surma samples, kohl and surma had significantly greater lead concentrations compared to kajal ($p = 0.015$, $p < 0.001$) (**Table 7**). Kohl and surma had medians of 170,000 ppm and 1,500 ppm respectively, while kajal had a median of 1.6 ppm. However, all three types had products with lead concentrations greater than 100,000 ppm.

Eyeliners production method

There was no significant difference in lead concentrations between manufactured eyeliners and homemade eyeliners (**Table 7**). There was no significant difference between traditional manufactured eyeliners and traditional homemade eyeliners (**Table 7**).

Eyeliners production region

Afghan-made eyeliners had significantly greater lead concentrations than products made in "Other Countries" (i.e., China, Egypt, Ethiopia, India, Morocco, Mexico, Pakistan, and Saudi Arabia) ($p = 0.016$) (**Table 7**). Both Afghan and Other Countries' eyeliners had significantly greater lead concentrations compared to US/Eu-made products ($p < 0.001$, $p = 0.0011$). The difference in lead concentrations is reflected in the medians of all three variables where Afghan products had a median of 29 ppm lead, Other Countries' products had a median of 2.8 ppm lead, and US/Eu products had a median of 0.94 ppm lead. Traditional Afghan-made products had significantly greater lead concentrations than traditional US/Eu-made products ($p = 0.01863$) (**Table 7**).

Eyeliners texture

Powdered eyeliners had significantly more lead than cream eyeliners ($p < 0.001$) (**Table 7**). 63.51% of cream eyeliners were <RL as opposed to 21.57% of powder eyeliners. The median for all powder

eyeliners was 690 ppm, whereas the median for cream eyeliners was 0.45 pm. Traditional powder eyeliners had significantly greater lead concentrations than traditional cream eyeliners ($p < 0.001$) (**Table 7**). The median for traditional powder eyeliners was 410,000 ppm, whereas the median for traditional cream eyeliners was 1.4 ppm.

Lead -free labels

There was no statistically significant difference between eyeliners with a “lead-free” label and eyeliners without (**Table 7**). Both categories had products with similar ranges.

Table 7: Summary of lead concentrations in eyeliners by eyeliner characteristics.

Eyeliner characteristics	Variables	n	% samples <RL	Median [range] (ppm)	p-value
<i>Eyeliners</i>	Traditional	80	20% (16)	10 [<0.2, 840,000]	Traditional > Nontraditional (<0.001)
	Nontraditional	63	76.67% (44)	0.06 [<0.0009, 1,800]	
<i>Type¹</i>	Kohl	26	30.77% (8)	170,000 [<0.00035, 840,000]	Kruskal-Wallis (0.0021) Kohl > Kajal (0.015) Surma > Kajal (<0.001)
	Kajal	27	25.92% (7)	1.6 [<0.0009, 130,000]	
	Surma	20	5% (1)	1,500 [<1.1, 670,000]	
<i>Production method (all)</i>	Manufactured	111	50.45% (56)	0.67 [<0.000701, 670,000]	No significant differences in lead concentrations
	Homemade ¹	9	0%	8.1 [0.9, 720,000]	
<i>(Traditional)</i>	Manufactured	51	24.41% (15)	3.2 [<0.0003, 670,000]	No significant difference in lead concentrations
	Homemade	9	0%	8.1 [0.9, 720,000]	
<i>Production country (all)</i>	US/Eu	35	65.71% (23)	0.94 [<0.0009, 600,000]	Kruskal-Wallis (<0.001) Other > US/Eu (0.0011) Afghanistan > US/Eu (<0.001) Afghanistan > Other (0.016)
	Afghanistan ¹	9	0%	29 [1.1, 630,000]	
	Other Countries	60	36.67% (22)	2.8 [<0.00071, 840,000]	
<i>(Traditional)</i>	US/Eu	10	20% (2)	1.5 [<0.0009, 600,000]	Afghanistan > US/Eu (0.019)
	Afghanistan	9	0%	29 [1.1, 630,000]	
	Other Countries	39	28.2% (11)	5.8 [<0.00071, 840,000]	
<i>Texture (all)</i>	Powder	51	21.57% (11)	690 [<0.0009, 840,000]	Powder > Cream (<0.001)
	Cream	74	63.51% (47)	0.45 [<0.003, 384.5]	
<i>(Traditional)</i>	Powder	33	0%	410,000 [0.5, 840,000]	Powder > Cream (<0.001)
	Cream	40	37.5% (15)	1.4 [<0.003, 384.5]	
<i>“Lead-free” label?²</i>	Lead-free label	11	0%	130,000 [0.5, 610,000]	No significant differences in lead concentrations
	No lead-free	31	0%	8.3 [0.4, 720,000]	

¹ All products within this category were traditional eyeliners.

Comparing current eyeliner testing methods

All eyeliners

For all eyeliners, the range for laboratory-derived lead concentrations was 0.4 ppm lead to 785,500 ppm lead, whereas the range for XRF RM Standard-derived concentrations was 3.2 ppm lead to 993,561 ppm lead (**Table 8**). The linear regression slope estimate was 1.43 with an R^2 value of 0.80 (95% CI [1.16, 1.71], $p < 0.001$) (**Figure 1**). The slope suggests that XRF RM Standard had a positive bias. The Pearson's and Spearman's coefficients were 0.89 and 0.86 respectively which indicates that there was a moderately strong linear relationship and moderately strong monotonic relationship between the laboratory-derived values and XRF RM Standard-derived values (**Figure 1**). Above 400,000 ppm, the XRF RM Standard measurements were between 900,000 ppm and 1×10^6 ppm (100% lead) despite none of the laboratory-derived lead concentrations being greater than 800,000 ppm lead. See **Appendix C** for more statistical data.

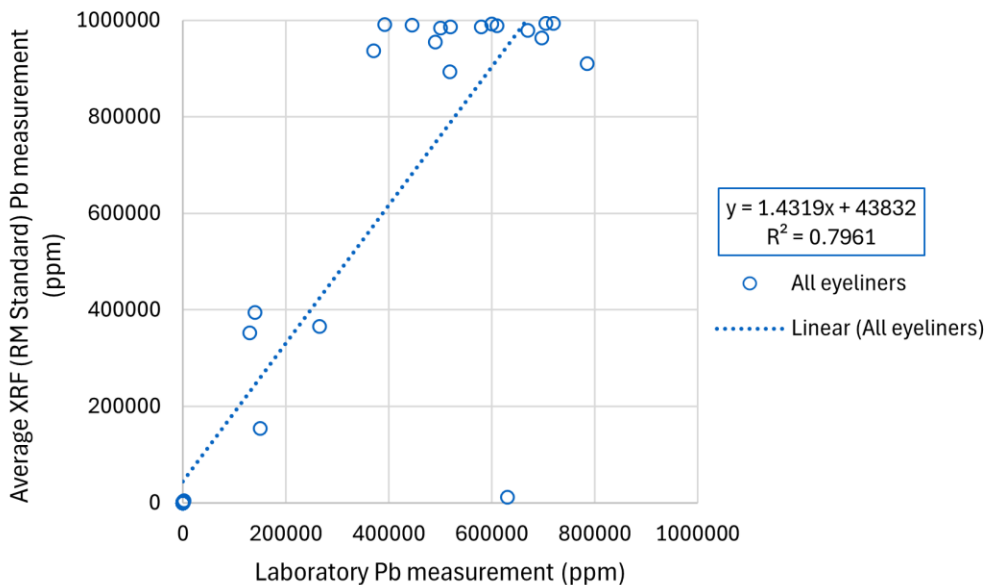


Figure 1: Linear regression to compare laboratory Pb measurements (ICP-MS, ICP-OES, GFAA) with average XRF Pb measurements for **all eyeliners** ($n=37$) (95% CI [1.16, 1.71], $p < 0.001$). Pearson's correlation coefficient = 0.89. Spearman's correlation coefficient = 0.86.

Eyeliners <400,000 ppm lead

For eyeliners <400,000 ppm, the range for laboratory-derived lead concentrations was 0.4 ppm lead to 391,500 ppm lead, whereas the range for XRF RM Standard-derived concentrations was 3.2 ppm lead to 990,619 ppm lead (**Table 8**). The slope of the regression line was about 2.3 with an R^2 value of 0.94 (95% CI [1.85, 2.74], $p < 0.001$) (**Figure 2**). XRF RM Standard had a positive bias and overestimated lead

concentrations in products. The Pearson’s and Spearman’s coefficients were 0.97 and 0.79 respectively which indicates a strong linear relationship and a moderate monotonic relationship between the laboratory-derived concentrations and XRF RM Standard-derived values (**Figure 2**). See **Appendix C** for more statistical data.

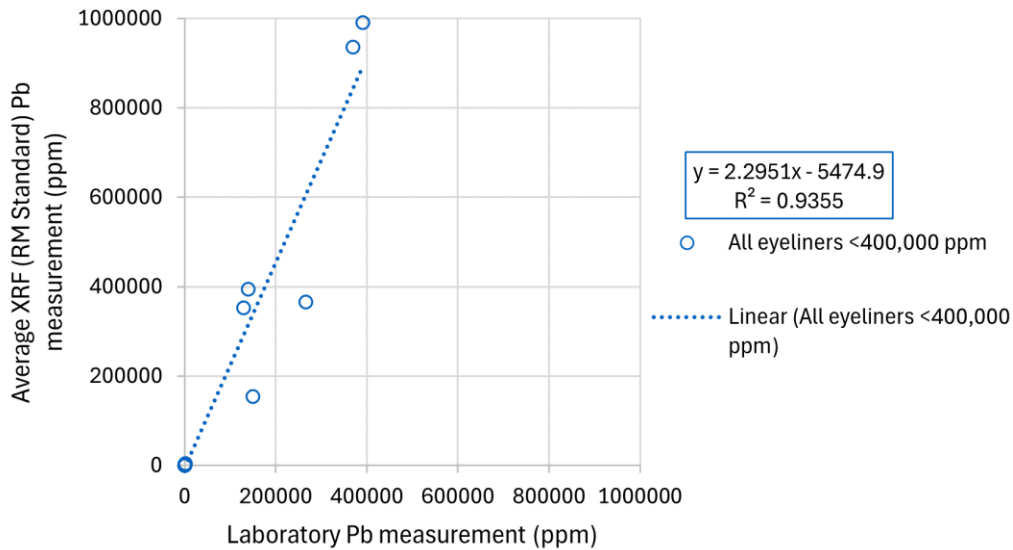


Figure 2: Linear regression to compare laboratory Pb measurements (ICP-MS, ICP-OES, GFAA) with average XRF Pb measurements for **eyeliners <400,000 ppm Pb** (n=23) (95% CI [1.85, 2.74], $p < 0.001$). Pearson’s correlation coefficient = 0.97. Spearman’s correlation coefficient = 0.79.

Powder eyeliners <400,000 ppm lead

Powder eyeliners <400,000 ppm lead had a range of 32 ppm lead to 391,500 ppm lead for laboratory-derived lead concentrations, and 31 ppm lead to 990,619 ppm lead for XRF RM Standard-derived lead concentrations (**Table 8**). The regression slope estimate was 2.3, and the R^2 value was 0.94 (95% CI [1.83, 2.86], $p < 0.001$) (**Figure 3**). The slope indicates that XRF RM Standard had a positive bias when measuring powdered eyeliners. The Pearson’s and Spearman’s coefficients were 0.95 and 0.74 respectively which indicates a strong linear relationship and a moderate monotonic relationship between the laboratory-derived concentrations and the XRF RM Standard-derived concentrations (**Figure 3**). See **Appendix C** for more statistical data.

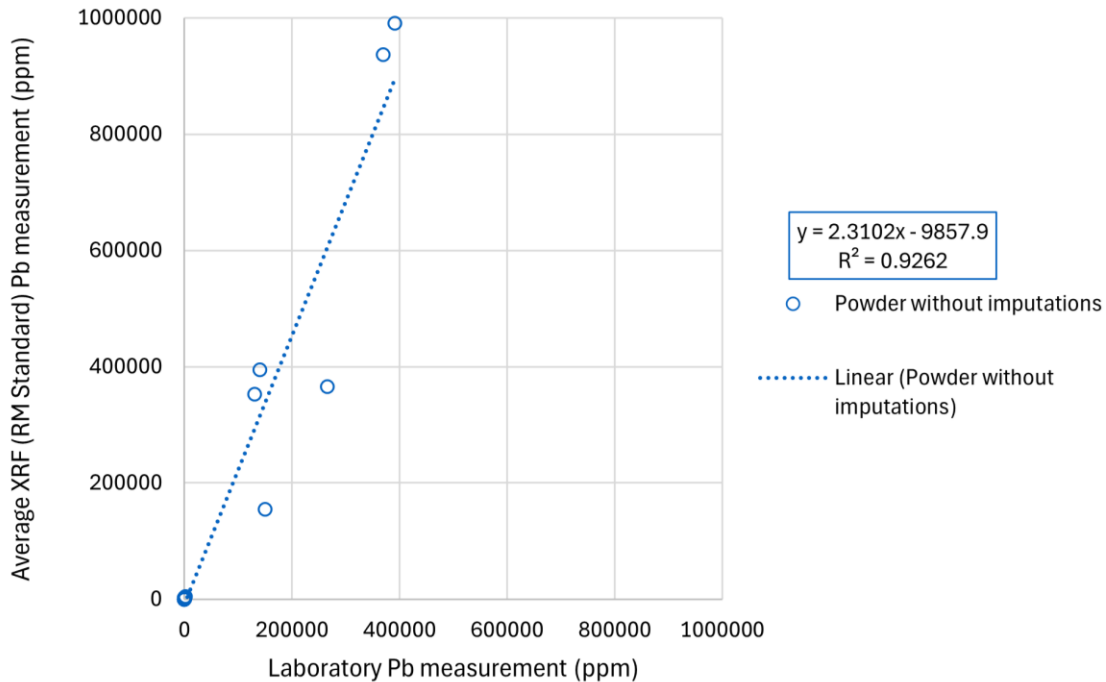


Figure 3: Linear regression to compare laboratory Pb measurements (ICP-MS, ICP-OES, GFAA) with average XRF Pb measurements for **powder eyeliners <400,000 ppm Pb** (n=15) (95% CI [1.83, 2.86], $p < 0.001$). Pearson's correlation coefficient = 0.95. Spearman's correlation coefficient = 0.74.

Cream eyeliners <90 ppm lead

For cream eyeliners <90 ppm lead, the range for laboratory-derived concentrations was 0.45 ppm to 29 ppm, whereas the range for XRF RM Standard-derived concentrations was 3.2 ppm to 38.6 ppm (**Table 8**). The slope of the regression line was 0.52 with an R^2 value of 0.13 (95% CI [-0.41, 1.44], $p = 0.22$) (**Figure 4**). The low R^2 value, along with the insignificant p-value and a 95% confidence interval that spans zero, indicates that there is not a significant linear relationship between the laboratory-derived lead concentrations and the XRF RM Standard-derived lead concentrations. The Pearson's and Spearman's coefficients were 0.36 and 0.40 respectively indicating a weak linear and a weak monotonic relationship between the laboratory-derived concentrations and the XRF RM Standard-derived concentrations (**Figure 4**). See **Appendix C** for more statistical data.

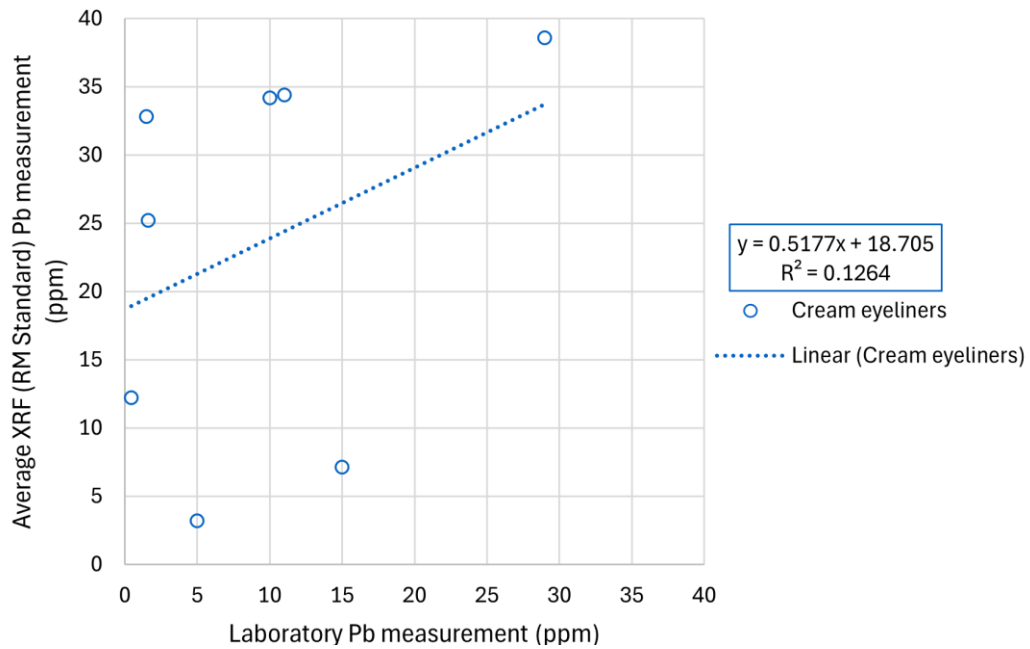


Figure 4: Linear regression to compare laboratory Pb measurements (ICP-MS, ICP-OES, GFAA) with average XRF Pb measurements for **cream eyeliners <90 ppm Pb** (n=8) (95% CI [-0.41, 1.44], p=0.22). Pearson's correlation coefficient = 0.36. Spearman's correlation coefficient = 0.40.

Table 8: Linear regression descriptive statistics for comparison between laboratory Pb measurements (ICP-MS, ICP-OES, GFAA) with average XRF (XRF RM Standard) Pb measurements for eyeliners (n=37).

Eyeliners analyzed	n	Method	Mean +/- SD (ppm)	Median [range] (ppm)
<i>All eyeliners</i>	37	Laboratory	268,110 +/- 289,634	140,000 [0.4, 785,500]
		XRF RM Standard	427,733 +/- 464,811	154,133 [3.2, 993,561]
<i>Eyeliners <400,000 ppm</i>	23	Laboratory	63,003 +/- 122,171	15 [0.4, 391,500]
		XRF RM Standard	139,122 +/- 289,903	34.4 [3.2, 990,619]
<i>Powder eyeliners <400,000 ppm</i>	10	Laboratory	144,898 +/- 15,2647	135,000 [32, 391,500]
		XRF RM Standard	319,706 +/- 37,5943	253,258 [31, 990,619]
<i>Cream eyeliners <90 ppm</i>	8	Laboratory	9.19 +/- 9.57	7.5 [0.45, 29]
		XRF RM Standard	23.46 +/- 13.93	29 [3.2, 38.6]

XRF method optimization

Calibrants

All calibrants

For all calibrants, the range was 5 ppm lead to 842,600 ppm lead. The slope for XRF RM Standard was 1.23 and the R^2 value was 0.99 (95% CI [1.17, 0.81], $p < 0.001$) (**Figure 5**). Pearson's and Spearman's coefficients were 0.99 indicating a strong linear relationship and a strong monotonic relationship between the actual calibrant lead concentrations and the XRF RM Standard-derived concentrations (**Figure 5**). The slope for XRF RM 180 was 0.78 and the R^2 value was 1.0 (95% CI [0.75, 0.81], $p < 0.001$) (**Figure 5**). Pearson's and Spearman's coefficients were both 1.0 indicating a strong linear relationship and a strong monotonic relationship between the actual calibrant lead concentrations and the XRF RM 180-derived concentrations (**Figure 5**). The slope for XRF Soil was 1.12 and the R^2 value was 0.99 (95% CI [1.05, 1.18], $p < 0.001$) (**Figure 5**). Pearson's and Spearman's coefficients were 0.99 indicating a strong linear relationship and a strong monotonic relationship between the actual calibrant lead

concentrations and the XRF Soil-derived concentrations (**Figure 5**). XRF Soil and XRF RM Standard both had a positive bias, whereas XRF RM 180 had a negative bias. See **Appendix D** for more statistical information.

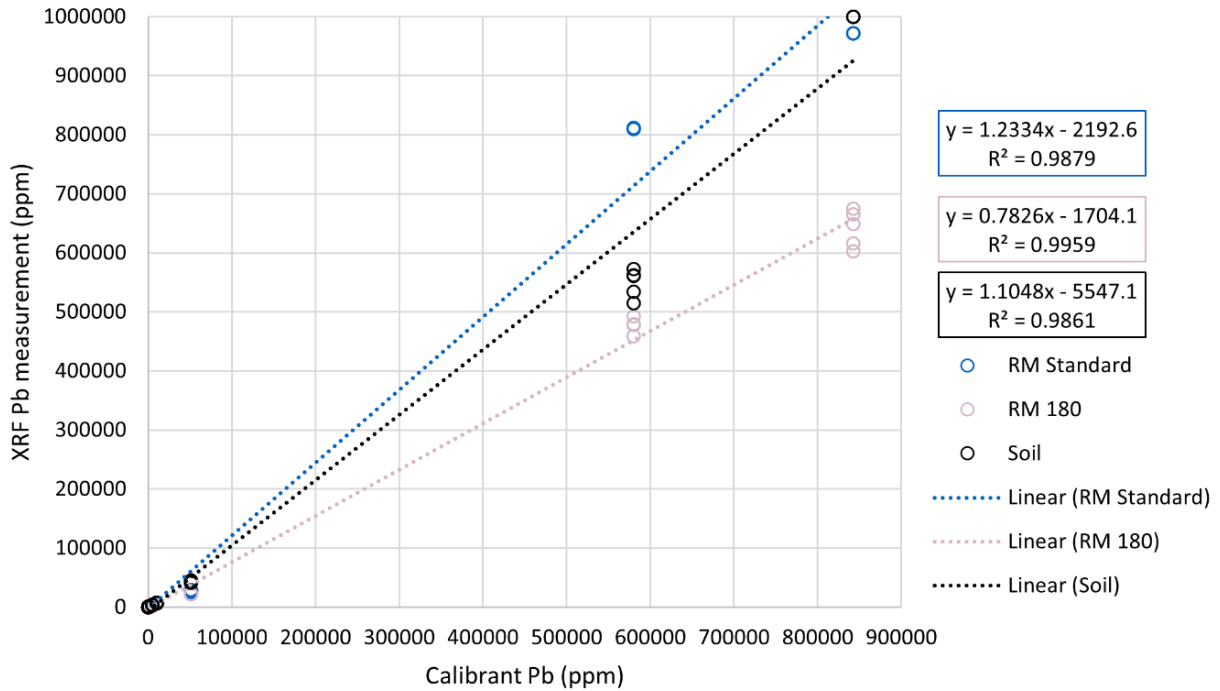


Figure 5: Linear regressions assessing the performance of the three XRF methods for **all calibrants** (5 ppm – 842,600 ppm, n=11). XRF RM Standard (95% CI [1.17, 0.81], p <0.001). Pearson’s correlation coefficient = 0.99. Spearman’s correlation coefficient = 0.99. XRF RM 180 (95% CI [0.75, 0.81], p <0.001). Pearson’s correlation coefficient = 1.0. Spearman’s correlation coefficient = 1.0. XRF Soil (95% CI [1.05, 1.18], p <0.001). Pearson’s correlation coefficient = 0.99. Spearman’s correlation coefficient = 0.99.

Calibrants <400,000 ppm lead

For calibrants <400,000 ppm lead, the range was 5 ppm lead to 51,300 ppm lead. The slope for XRF RM Standard was 0.53 and the R² value was 1.0 (95% CI [0.51, 0.56], p <0.001) (**Figure 6**). Pearson’s and Spearman’s coefficients were 1.0 and 0.99 respectively, indicating a strong linear relationship and a strong monotonic relationship between the actual calibrant lead concentrations and the XRF RM Standard-derived concentrations (**Figure 6**). The slope for XRF RM 180 was 0.42 and the R² value was 1.0 (95% CI [0.42, 0.84], p <0.001) (**Figure 6**). Pearson’s and Spearman’s coefficients were 1.0 and 0.99 respectively, indicating a strong linear relationship and a strong monotonic relationship between the actual calibrant lead concentrations and the XRF RM 180-derived concentrations (**Figure 6**). The slope for XRF Soil was 0.87 and the R² value was 1.0 (95% CI [0.82, 0.87], p <0.001) (**Figure 6**). Pearson’s and Spearman’s coefficients were 1.0 and 0.99 respectively, indicating a strong linear relationship and a

strong monotonic relationship between the actual calibrant lead concentrations and the XRF Soil-derived concentrations (**Figure 6**). XRF Soil and XRF RM Standard had a positive bias, and XRF RM 180 had a negative bias. XRF Soil's slope was the closest to 1.0 out of the three methods. See **Appendix D** for more statistical information.

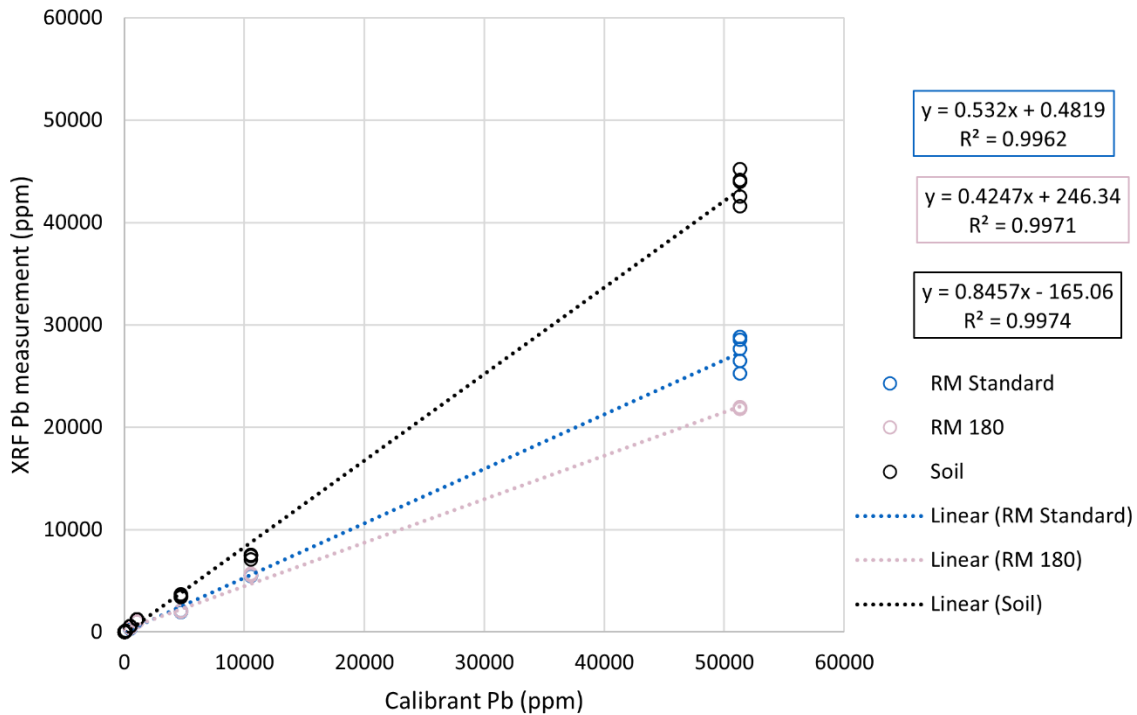


Figure 6: Linear regressions assessing the performance of the three XRF methods for **calibrants <400,000 ppb Pb** (5 ppm – 51,300 ppm, n=9). XRF RM Standard (95% CI [0.51, 0.56], p <0.001). Pearson's correlation coefficient = 1.0. Spearman's correlation coefficient = 0.99. XRF RM 180 (95% CI [0.42, 0.84], p <0.001). Pearson's correlation coefficient = 1.0. Spearman's correlation coefficient = 0.99. XRF Soil (95% CI [0.82, 0.87], p <0.001). Pearson's correlation coefficient = 1.0. Spearman's correlation coefficient = 0.99. See Table C2 for more statistical data.

Eyeliners

All tested eyeliners

The laboratory-derived concentrations for the subset of eyeliners had a range of 32 ppm lead to 720,000 ppm lead. The slope for XRF RM Standard was 1.54 and the R² value was 0.85 (95% CI [1.32, 1.61], p <0.001) (**Figure 7**). Pearson's and Spearman's coefficients were 0.92 and 0.96 respectively, indicating a strong linear relationship and a strong monotonic relationship between the laboratory-derived lead concentrations and the XRF RM Standard measured concentrations (**Figure 7**). The slope for XRF RM 180 was 1.38 and the R² value was 0.61 (95% CI [1.18, 1.57], p <0.001) (**Figure 7**). Pearson's and Spearman's coefficients were 0.78 and 0.87 respectively, indicating a moderate linear relationship and a moderate monotonic relationship between the laboratory-derived lead concentrations and the XRF RM 180

measured concentrations (**Figure 7**) The slope for XRF Soil was 1.62 and the R^2 value was 0.86 (95% CI [1.43, 1.66], $p < 0.001$) (**Figure 7**). Pearson's and Spearman's coefficients were 0.98 and 0.93 respectively, indicating a strong linear relationship and a strong monotonic relationship between the laboratory-derived lead concentrations and the XRF Soil-derived concentrations (**Figure 7**). All three methods had a positive bias and overestimated the lead concentration in the eyeliners, and XRF RM 180 had a slope closest to 1.0. For eyeliners above 400,000 ppm lead, all three XRF method measurements were between 900,000 ppm and 1×10^6 ppm (100% lead) despite none of the laboratory-derived lead concentrations being above 800,000 ppm lead. See **Appendix D** for more statistical information.

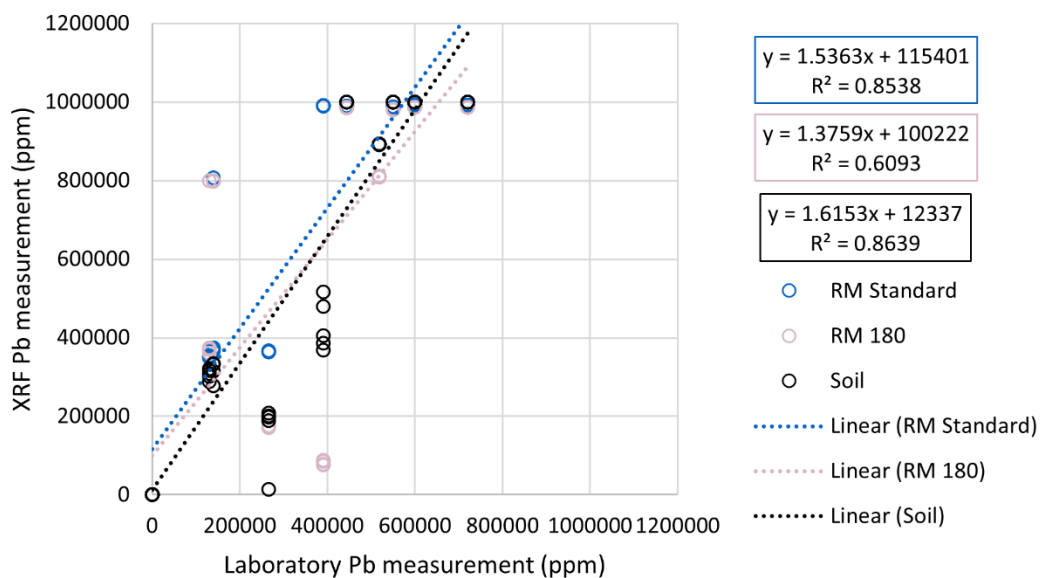


Figure 7: Linear regressions assessing the performance of the three XRF methods by comparing laboratory measurements with XRF methods for *all tested eyeliners* (GFAA (n=8), ICP-MS (n=4), 32 ppm - 720,000 ppm). XRF RM Standard (95% CI [1.32, 1.61], $p < 0.001$). Pearson's correlation coefficient = 0.92. Spearman's correlation coefficient = 0.96. XRF RM 180 (95% CI [1.18, 1.57], $p < 0.001$). Pearson's correlation coefficient = 0.78. Spearman's correlation coefficient = 0.87. XRF Soil (95% CI [1.43, 1.66], $p < 0.001$). Pearson's correlation coefficient = 0.98. Spearman's correlation coefficient = 0.93.

Tested eyeliners <400,000 ppm lead

For eyeliners <400,000 ppm lead, the range was 32 ppm to 391,500 ppm. The slope for XRF RM Standard was 2.22 and the R^2 value was 0.83 (95% CI [1.32, 1.61], $p < 0.001$) (**Figure 8**). Pearson's and Spearman's coefficients were 0.91 and 0.94 respectively, indicating a strong linear relationship and a strong monotonic relationship between the laboratory-derived lead concentrations and the XRF RM Standard measured concentrations (**Figure 8**). The slope for XRF RM 180 was 0.45 and the R^2 value was 0.0005 (95% CI [-0.39, 0.68], $p = 0.58$), which suggests there was not a significant relationship between the XRF

RM 180 derived lead concentrations and the laboratory-derived concentrations (**Figure 8**). The slope for XRF Soil was 0.922 and the R^2 value was 0.59 (95% CI [0.71, 1.14], $p < 0.001$) (**Figure 8**). Pearson's and Spearman's coefficients were 0.77 and 0.80 respectively, indicating a moderate linear relationship and a moderate monotonic relationship between the laboratory-derived lead concentrations and the XRF Soil-derived concentrations (**Figure 8**). See **Appendix D** for more statistical information.

XRF RM Standard had a slope of 2.3 for both the larger set of 37 eyeliner and for the subset of 12 eyeliner. Having similar slopes suggests that the subset of 12 eyeliner is a good representation of eyeliner found in the field despite its small sample size. XRF RM Standard was positively biased, whereas XRF RM 180 and XRF Soil were negatively biased. XRF Soil had a slope closest to 1.0, however, it had a moderately low R^2 value, and moderate Pearson and Spearman correlation coefficients compared to XRF RM Standard.

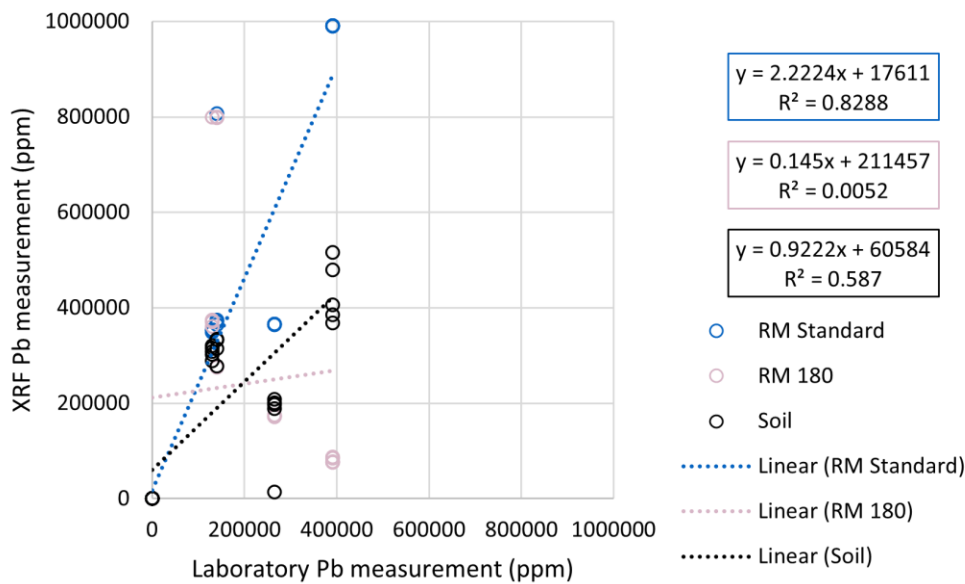


Figure 8: Linear regressions assessing the performance of the three XRF methods by comparing laboratory measurements with XRF methods for **eyeliners <400,000 ppm** (GFAA (n=4), ICP-MS (n=2), 32 ppm - 391,500 ppm). XRF RM Standard (95% CI [1.32, 1.61], $p < 0.001$). Pearson's correlation coefficient = 0.91. Spearman's correlation coefficient = 0.94. XRF RM 180 (95% CI [-0.39, 0.68], $p = 0.58$). Pearson's correlation coefficient = 0.72. Spearman's correlation coefficient = 0.46. XRF Soil (95% CI [0.71, 1.14], $p < 0.001$). Pearson's correlation coefficient = 0.77. Spearman's correlation coefficient = 0.80.

Discussion

Through our investigation we were able to answer our three aims. Our findings suggest that there are certain characteristics in traditional and nontraditional eyeliners that impact lead concentration and the XRF's ability to measure lead concentration.

Eyeliner characteristics

Our results indicated that traditional eyeliners had significantly greater lead concentrations compared to nontraditional eyeliners (**Table 7**). This may be due to ingredient differences. Two studies investigated the main components of traditional eyeliners using Scanning Electron Microscopy (SEM) (**Table 10**).^(21,22) Hardy et al. (2002) found that the main component of many traditional powder eyeliners was galena followed by zincite (ZnO).⁽²²⁾ Filella et al. (2020) found that many eyeliners with >100,00 ppm lead were made of PbS, which we know is the composition of galena.⁽²¹⁾ In our dataset, 29 eyeliners had >100,000 ppm lead, and all were traditional powder eyeliners. Their listed ingredients included “kohl powder”, “surma black rock”, and “natural surma” which may be terms for galena. Two products were >100,000 ppm but listed their main ingredient as “kohl powder from antimony” and “stibnite” which are other names for Sb₂S₃, a lead-free mineral that is more expensive than galena. Although labeling suggested the eyeliners were lead-free, they had >100,000 ppm lead according to laboratory-derived results. Other ingredients listed for traditional powder eyeliners included camphor (bhimseni), frankincense gum, powdered rose extract, Sankh (translation unknown), and hareer (an herb).

Cream traditional eyeliners had significantly less lead than powder traditional eyeliners (the maximum was 384.5 ppm lead compared to 840,000 ppm lead), but some products still violated U.S. regulations. Ingredients listed on product labels included dimethicone (a silicone), beeswax, and animal fats such as ghee. One product was made with stone surma and water. Filella et al.'s SEM-EDXS analysis found that most traditional cream eyeliners they analyzed were composed of carbon, and many contained a silicone compound.⁽²¹⁾ Hardy et al. found that traditional cream eyeliners were mainly composed of zincite, amorphous carbon, and paraffin wax.⁽²²⁾ None of the cream eyeliners Hardy et al. tested contained lead. Navarro-Tapia et al. (2021) and Gouitaa et al. (2016), also found that traditional cream eyeliners (kohl pastes and kohl pencils) contained lower lead concentrations than traditional powder eyeliners (kohl powder).^(17,56) Despite having less lead than powder eyeliners, our study found that cream eyeliners, particularly traditional ones, may still be contaminated with lead.

Kohl and surma had significantly greater lead concentrations than kajal (**Table 7**). No other studies were found that differentiated between the three types of eyeliners and their lead content. Kohl is often a powder made from galena, and therefore is more likely to have lead content above 100,000 ppm (**Table 1**).^(28,37) Surma tends to be made with galena as well; however, community members informed us that there are two types of surma, one made from grinding minerals- likely galena- and the other from mixing soot and animal fat (**Table 1**).^(12,32,33) When we looked at the descriptive statistics for surma, we saw that the range reflects this difference in type as the minimum is less than 1.1 ppm whereas the maximum is 670,000 ppm (**Table 7**). This could be an example where surma made from soot and oil had lower lead concentrations than surma made from ground minerals.

Kajal, on the other hand, had significantly lower lead concentrations. According to the literature, kajal is often made using burnt organic matter mixed with animal fat or vegetable oil (**Table 1**).^(18,30,37) Unfortunately, despite having lower lead concentrations than kohl and surma, there were still kajal products with lead concentrations above 100,000 ppm, which suggests consumers cannot assume one traditional eyeliner is safer than another based on type alone (**Table 7**).

U.S. and European-made eyeliners had significantly less lead than eyeliners made in Afghanistan and Other Countries (**Table 7**). Afghanistan has no regulations on lead in cosmetics; therefore, it is unsurprising that Afghan-made products also had the highest lead levels in our dataset (**Table 2**). McMichael and Stoff (2017) found comparable results when they looked at eyeliners made in Afghanistan which had lead ranging from 400 ppm to over 800,000 ppm (**Table 9**).⁽³³⁾

Other Countries consisted of eyeliners made in India, Pakistan, China, Mexico, Saudi Arabia, Ethiopia, Egypt, and Morocco. India and Saudi Arabia have regulations that limit the amount of lead in cosmetics to 20 ppm and 10 ppm respectively (**Table 2**).^(53,54) Despite regulations, we found eyeliners made in India and Saudi Arabia with more than 100,000 ppm lead, which suggests a major ingredient was lead. Pakistan has no regulations on lead in cosmetics, and we tested products made in Pakistan that had more than 100,000 ppm lead (**Table 2**).

The finding of eyeliners with high lead concentrations made in countries that have cosmetic regulations suggests these regions have less oversight in eyeliner production. When we looked at eyeliners with “lead-free” labeling, we found six products with lead-free labels that contained over 100,000 ppm lead. All except one were made outside of the United States or Europe. The sixth was made in Morocco and

Spain, but it was unclear which ingredients were from Spain and which were from Morocco. Filella et al. (2020) and Daar et al. (2017) found similar results when they tested several traditional eyeliners with “lead-free” labeling (**Table 9**).^(21,23) They tested eyeliners that claimed to be lead-free but contained more than 1,000 ppm lead. The misleading labels in our study and previous studies indicate that consumers must be careful not to rely on labeling when purchasing traditional eyeliners. It also suggests that eyeliners produced in certain LMICs may not be enforcing labels as strictly as the United States and Europe.

Out of all traditional US- and European-made eyeliners, two eyeliners were above the FDA limit of 10 ppm. One eyeliner produced in the United States and purchased by us on Etsy had 600,000 ppm lead; however, when we went back to check for this product on Etsy, the shop and item were unavailable. The second eyeliner had 11 ppm lead; however, the sample sent to the lab was below the recommended 1 g which may impact the accuracy of the lead measurement.

Comparing current eyeliner testing methods

Based on our results, it appears that XRF RM Standard is an appropriate method for cosmetic screening because it had an $R^2 > 0.70$ and moderate to strong Pearson's and Spearman's correlation coefficients. We found that XRF RM Standard was oversaturated above 400,000 ppm lead, resulting in a curvilinear relationship. For screening purposes, oversaturation may not pose a problem because it still informs community members that their eyeliner is hazardous, and they should no longer use it. We were unable to ascertain whether XRF RM Standard has a maximum LOD as that may have influenced the instrument's ability to detect lead in eyeliners with high lead concentrations.

One finding of note was that cream eyeliners did not have a significant linear relationship between the laboratory-derived lead concentrations and the XRF RM Standard-derived concentrations (**Figure 4**). This could be due to a few factors. The first is that cream eyeliners had lead concentrations closer to the XRF RM Standard's mean LOD of 11.1 ppm, as opposed to powder eyeliners which had higher lead concentrations. Sample weight may also play a role, as cream products, especially eye pencils, had fewer grams per sample compared to powder products, and therefore may not have provided the thickness of $\frac{3}{4}$ in that is optimal for XRF testing. Finally, the XRF may be better suited to dry samples compared to liquid or cream samples, which means that XRF RM Standard is not the best method for screening cream eyeliners.⁽⁶⁶⁾ PHSKC investigators should continue to send cream eyeliners for laboratory chemical analysis rather than XRF screening.

Daar et al. (2017) used a non-portable XRF to measure lead in traditional eyeliners (**Table 10**).⁽²³⁾ They only analyzed powder eyeliners with high lead concentrations (30,000 - 570,000 ppm) and therefore did not run into the issues we did with eyeliners having lead concentrations close to or below the instrument's LOD. Filella et al. (2020) used a portable XRF on powder and cream traditional eyeliners (**Table 10**).⁽²¹⁾ Five of the seven cream products did not have lead according to the XRF, which may be due to the cream products falling close to the XRF's LOD of <20 mg/kg, which is similar to what we found with our cream eyeliners. Filella et al. did not verify their XRF measurements with separate laboratory chemical analyses as we did.

Our findings as well as findings from other studies suggest that the portable XRF is a powerful tool for screening powder eyeliners with hazardous lead concentrations, but it is not well equipped for cream eyeliners. We suggest that PHSKC investigators continue to use XRF RM Standard in the field for powder eyeliners, but that cream eyeliners be sent for laboratory chemical analysis.

XRF RM Standard had a positive bias, which makes it more protective for communities. From a regulatory standpoint, a positive bias may concern brands and manufacturers as it could result in false positive readings for products. To prevent false positives, samples should be sent for laboratory analysis in tandem with XRF analysis, especially if they are close to the XRF's LOD.

XRF method optimization

We found that XRF Soil was the most optimal method for measuring lead in soil when testing the calibrants because it not only had an R^2 value >0.70, and strong Pearson's and Spearman's correlation coefficients, but it also had a slope closest to 1.0.

We found one study using calibrants to improve their XRF analysis method for lead in eyeliners. Daar et al. (2017) developed in-house standards by mixing known concentrations of iron, zinc, and lead with graphite (**Table 10**).⁽²³⁾ Their minimum detectable limit for their XRF was 0.020% or 2000 ppm lead, which is much higher than the LODs for our three XRF methods. Similar to our study, they found that error terms for the XRF measurements increased as lead concentrations in the eyeliners decreased.

When testing eyeliners, XRF Soil had a slope of 0.92, which was the closest to 1.0 out of all three XRF methods; however, its R^2 value did not meet the EPA criteria of 0.70 (**Figure 8**). With a larger sample size and more robust validation, we may be able to improve XRF Soil's R^2 value.

Limitations and Future Studies

When we investigated eyeliner characteristics, we used three different laboratory chemical analysis methods to analyze eyeliners – ICP-MS, ICP-OES, and GFAA. All three had varying RLs, with ICP-OES having the highest RL and thus needing the most imputations. For future studies, eyeliners should be sent for ICP-MS or GFAA testing due to their lower RLs.

Another limitation we had was the age of our XRF analyzer. XRF performance declines as it gets older. The XRF we used was six years old, and its last calibration occurred in 2021.

When determining an optimal XRF method for eyeliner testing, we were limited by our small subset of 11 calibrants and 12 eyeliners. We acknowledge that it would have been ideal if we had tested all eyeliners using the three methods. Future studies could investigate the three methods more thoroughly with a larger sample. Future studies could also look at developing an entirely new XRF method that does not rely on the instrument's built-in automatic calibration setting.

Table 9. Studies examining metal concentrations *in traditional eyeliners*.

Study	Country	Cosmetic term used	Range Pb (ppm)	Analysis method	Community engagement?
Current study	United States	Kohl, kajal, surma	<0.2 - 840,000	ICP-MS ^a ; ICP-OES ^b ; GFAA ^c	Yes- items donated by community members
New York City Department of Health & Mental Hygiene (2011-present) ⁽⁵⁷⁾	United States	Kohl, kajal, surma	<0.33 - 980,000	Laboratory analysis ^d	Yes - via investigations of lead poisonings and store surveys
King County Lead & Toxics- 2022 Product testing events ⁽⁵⁸⁾	United States	Kohl, kajal, surma	390,000 - 840,000	Laboratory analysis ^d	Yes - items donated by community members
Navarro-Tapia et al. 2021 ⁽¹⁷⁾	Spain	Kohl-based eye cosmetics	1.73 - 410, 806	ICP-OES	No
Rasheed et al. 2021 ⁽²⁶⁾	Iraq	Kohl	0.727 - 1,491	ICP-OES	No
Buksh et al. 2020 ⁽²⁷⁾	Pakistan	Kohl	<2.32 - 199.9	FAAS ^e	No
Filella et al. 2020 ⁽²¹⁾	Switzerland	Kohl	<20 - 467,632	XRF ^f ; SEM-EDXS ^g	No
Daar et al. 2017 ⁽²³⁾	Jordan	Kohl	30,000 - 570,000	XRF	No
McMichael & Stoff 2017 ⁽³³⁾	Afghanistan	Surma	400 - 830,000	ICP-MS	No
Gouitaa et al. 2015 ⁽⁵⁶⁾	Morocco	Kohl	1 - 973,800	AAS ^h	No
Hardy et al. 2002 ⁽²²⁾	United Arab Emirates	Kohl	<50,000 - 970,000	X-ray microanalyzer; X-ray diffraction	One item donated by student

^a Inductively coupled plasma mass spectrometry (ICP-MS)

^b Inductively coupled plasma optical emission spectroscopy (ICP-OES)

^c Graphite furnace atomic absorption spectroscopy (GFAA)

^d Method of analysis is generalized since both departments utilized a variety of measuring techniques.

^e Flame atomic absorption spectrometry (FAAS)

^f X-ray fluorescence (XRF)

^g Scanning electron microscopy with energy dispersive X-ray spectroscopy

^h Atomic absorption spectrometry (AAS)

Table 10. Studies examining metal concentrations in *non-traditional eye cosmetics*.

Study	Country	Cosmetic type	Range Pb (ppm)	Analysis method	Community engagement?
Current study	United States	Eyeliners, Eyeshadows, Mascaras	<0.0009 - 1,800	ICP-MS ^a ; ICP-OES ^b ; GFAA ^c	Yes- items donated by community members
New York City Department of Health & Mental Hygiene (2011-present) ⁽⁵⁷⁾	United States	Eyeliners, Eyeshadows	<0.36 - 930,000 ^d	Laboratory analysis ^e	Yes - via investigations of lead poisonings and store surveys
King County Lead & Toxics- 2022 Product testing events ⁽⁵⁸⁾	United States	Eyeliners, Eyeshadows	15 - 100	Laboratory analysis ^e	Yes – items donated by community members
Pawlaczyk et al. 2021 ⁽⁶¹⁾	Poland	Eyeshadows	0.0957 - 15.95	ICP-MS	No
Omoloye et al. 2020 ⁽⁷³⁾	Nigeria	Eyeshadows	<0.5 - 46.67	FAAS	No
Lim et al. 2018 ⁽⁷⁴⁾	South Korea	Eyeliners	<0.02 - 9.71	ICP-MS	No
Omenka & Adeyi 2016 ⁽⁷⁵⁾	Nigeria	Eyeliners	3.71 - 27.5	AAS ^f	No
Volpe et al. 2012 ⁽⁶²⁾	Italy	Eyeshadows	Chinese brands: 9.53 - 65.6 U.S brands: 0.25 - 2.75 Italian brands: 0.57 - 7.64	FAAS ^g	No
Charter et al. 2011 ⁽⁵⁹⁾	Canada	Eyeliners, Eyeshadows	<0.04 - 110	ICP-OES; ICP-MS	Yes - Asked community members to share the products they use daily
Al- Saleh et al. 2009 ⁽⁶³⁾	Saudi Arabia	Eyeshadows	0.42 - 58.72	ZAAS coupled to GTA ^h	No
Sainio et al. 2000 ⁽⁶⁰⁾	Finland	Eyeshadows	<0.5 - 16.8	EAAS ⁱ	No

^a Inductively coupled plasma mass spectrometry (ICP-MS)

^b Inductively coupled plasma optical emission spectroscopy (ICP-OES)

^c Graphite furnace atomic absorption spectroscopy (GFAA)

^d All but one product had Pb levels <36 ppm.

^e Method of analysis is generalized since both departments utilized a variety of laboratory techniques.

^f Atomic absorption spectrometry (AAS)

^g Flame atomic absorption spectrometry (FAAS)

^h Zeeman atomic absorption spectrophotometer coupled to graphite tube atomizer (ZAAS coupled to GTA)

ⁱ Electrothermal atomic absorption spectroscopy (EAAS)

Comparing current eyeliner testing methods

Based on our results, it appears that XRF RM Standard is an appropriate method for cosmetic screening because it had an $R^2 > 0.70$ and moderate to strong Pearson's and Spearman's correlation coefficients. We found that XRF RM Standard was oversaturated above 400,000 ppm lead, resulting in a curvilinear relationship. For screening purposes, oversaturation may not pose a problem because it still informs community members that their eyeliner is hazardous, and they should no longer use it. We were unable to ascertain whether XRF RM Standard has a maximum LOD for lead.

One finding of note was that cream eyeliners did not have a significant linear relationship between the laboratory-derived lead concentrations and the XRF RM Standard-derived concentrations (**Figure 4**). This could be due to a few factors. The first is that cream eyeliners had lead concentrations closer to the XRF RM Standard's mean LOD of 11.1 ppm, as opposed to powder eyeliners which had higher lead concentrations. Sample weight may also play a role, as cream products, especially eye pencils, had fewer grams per sample compared to powder products, and therefore may not have provided the thickness of $\frac{3}{4}$ inches that is optimal for XRF testing. Finally, the XRF may be better suited to dry samples compared to liquid or cream samples, which means that XRF RM Standard is not the best method for screening cream eyeliners.⁽⁶⁶⁾ PHSKC investigators should send cream eyeliners for laboratory chemical analysis rather than XRF screening.

Daar et al. (2017) used a non-portable XRF to measure lead in traditional eyeliners (**Table 10**).⁽²³⁾ They only analyzed powder eyeliners with high lead concentrations (30,000 - 570,000 ppm) and therefore did not run into the issues we did with eyeliners having lead concentrations close to or below the instrument's LOD. Filella et al. (2020) used a portable XRF on powder and cream traditional eyeliners (**Table 10**).⁽²¹⁾ Five of the seven cream products did not have lead according to the XRF, which may be due to the cream products falling close to the XRF's LOD of <20 mg/kg, which is similar to what we found with our cream eyeliners. Filella et al. did not verify their XRF measurements with separate laboratory chemical analyses as we did.

Our findings as well as findings from other studies suggest that the portable XRF is a powerful tool for screening powder eyeliners with hazardous lead concentrations, but it is not well equipped for cream eyeliners. We suggest that PHSKC investigators continue to use XRF RM Standard in the field for powder eyeliners, but that cream eyeliners be sent for laboratory chemical analysis.

XRF RM Standard had a positive bias, which makes it is more protective for communities. From a regulatory standpoint, a positive bias may concern brands and manufacturers as it could result in false positive readings for products. To prevent false positives, samples should be sent for laboratory analysis in tandem with XRF analysis, especially if they are close to the XRF's LOD.

XRF method optimization

We found that XRF Soil was the most optimal method for measuring lead in soil when testing the calibrants because it not only had an R^2 value >0.70 , and strong Pearson's and Spearman's correlation coefficients, but it also had a slope closest to 1.0.

We found one study using calibrants to improve their XRF analysis method for lead in eyeliners. Daar et al. (2017) developed in-house standards by mixing known concentrations of iron, zinc, and lead with graphite (**Table 10**).⁽²³⁾ Their minimum detectable limit for their XRF was 0.020% or 2000 ppm lead, which is much higher than the LODs for our three XRF methods. Similar to our study, they found that error terms for the XRF measurements increased as lead concentrations in the eyeliners decreased.

When testing eyeliners, XRF Soil had a slope of 0.92, which was the closest to 1.0 out of all three XRF methods; however, its R^2 value did not meet the EPA criteria of 0.70 (**Figure 8**). With a larger sample size and more robust validation, we may be able to improve XRF Soil's R^2 value.

Limitations and Future Studies

When we investigated eyeliner characteristics, we used three different laboratory chemical analysis methods to analyze eyeliners – ICP-MS, ICP-OES, and GFAA. All three had varying RLs, with ICP-OES having the highest RL and thus needing the most s. For future studies, eyeliners should be sent for ICP-MS or GFAA testing due to their lower RLs.

Another limitation we had was the age of our XRF analyzer. XRF performance declines as it gets older. The XRF we used was six years old, and its last calibration occurred in 2021.

When determining an optimal XRF method for eyeliner testing, we were limited by our small subset of 11 calibrants and 12 eyeliners. We acknowledge that it would have been ideal if we had tested all eyeliners using the three methods. Future studies could investigate the three methods more thoroughly with a larger sample. Future studies could also look at developing an entirely new XRF method that does not rely on the instrument's built-in automatic calibration setting.

Conclusion

Traditional eyeliners, particularly traditional powder eyeliners, are more likely to be made from galena and expose users to high lead concentrations as a result. Eyeliners made outside of the United States and Europe also had higher lead concentrations which may be due to a lack of regulation or a lack of enforcement. When we looked at traditional eyeliners with “lead-free” labels, we found some eyeliners made outside of the United States that claimed to be lead-free when they had over 100,000 ppm lead. Therefore, we recommend that consumers purchase nontraditional eyeliners and/or eyeliners made in the United States or Europe to limit lead exposure.

During field testing of cosmetics, we found that XRF RM Standard was good at screening powder eyeliners; however, it struggled with wetter eyeliners that had lead concentrations close to the LOD. Previous studies measuring lead in eyeliners used dry samples with higher lead concentrations. PHSKC investigators should keep the limitations of XRF RM Standard in mind when measuring, not just eyeliners, but other products in the field that may be similar to eyeliners such as soils, food, and other cosmetics. With more robust validation, XRF Soil could be a better method than XRF RM Standard. While the XRF is a powerful screening tool, cosmetics should continue to be sent for chemical analyses using the gold standard methods when determining whether a product, brand, or manufacturer is following lead regulations. We hope our findings provide community members with more guidance on safer practices for purchasing and using eyeliners, and provide public health investigators with a better understanding of the XRF's strengths and limitations for screening eyeliners.

Appendix

Appendix A: Sample preparation

For eyeliners purchased by the Research team, we bought at least two of every eyeliner. Both products were scanned with the XRF, and one was sent to the laboratory for further analysis. For eyeliners that contained less product, such as cream eye pencils, we purchased three or more of each. We used a spatula (LevGo 17211 smartSpatula Disposable Polypropylene Laboratory Spatula, 210 mm (Pack of 300), Blue) to remove product from cream eye pencils. New spatulas were used for each product to prevent cross-contamination. If the pencil was made of wood, we split it so that the product could be scraped out from within the pencil before being transferred to the plastic sample bag (Plymor Zipper Reclosable Plastic Bags, 2 Mil, 2" x 3" (Pack of 100): Industrial & Scientific). Sample bags were purchased so that their thickness did not interfere with the attenuation of the XRF's signal as recommended by Bruker.⁽⁶⁶⁾ Cream product was collected in one corner of the sample bag to increase the thickness of the sample. The sample should be at least $\frac{3}{4}$ inches (in) thick when analyzed through a plastic bag to obtain accurate results.⁽⁶⁶⁾ For cream eyeliners that were not pencils, we scraped product out of the container so that it was approximately $\frac{3}{4}$ in thick.

Powder eyeliners were gently scraped from their original container and into the sample bags using the spatula. Sample bags were lightly tapped to collect the powder in one area to maximize sample thickness.

Appendix B: Eyeliner characteristics

Table B1. Descriptive statistics to understand the central tendencies of different lab techniques for all eyeliners assessed.

Lab technique	Sample size	Reporting limit (RL) (ppm)	% sample s <RL	% samples ND	Min	1st quartile (ppm)	Median [range] (ppm)	3rd quartile (ppm)	Max (ppm)	Mean +/- SD (ppm)
GFAA ¹	66	0.30	12% (8)	3% (2)	<0.0003	0.8	5.0 [<0.003 - 720,000]	170	720,000	95,188 +/- 208,289
ICP-MS ^{2*}	11	0.35	0%	0%	10	241	391,500 [10-785,500]	608,000	785,500	17,658 +/- 118
ICP-OES ³	68	14.75	77.94% (53)	47% (32)	<0.0009	0.47	0.04 [<0.0009 - 840,000]	12.2	840,000	69,891 +/- 203,304
All methods	145	NA	42.1% (61)	22.1% (32)	<0.0009	2.8	3.2 [<0.0009 - 840,000]	113.1	840,000	102,370 +/- 225,007

1- GFAA measurements were conducted by NVL Laboratories. Each sample had a different RL due to differences in weight and matrix. We averaged all GFAA samples that were <RL to produce an average RL of 0.3. The range for <RL was [<0.16, <1.1]

2- ICP-MS measurements were conducted by the Environmental Health Laboratory and Trace Organics Analysis Center. The reporting limit was calculated by taking the lowest calibrant used on the instrument (35 ng) and dividing it by a representative sample aliquot mass (100mg).

3- ICP-OES measurements were conducted by NVL Laboratories. Each sample had a different reporting limit due to differences in weight and matrix. We averaged all ICP-OES samples below the reporting limit to produce an average reporting limit. The range of <RL was [<3.4, <160]

*No imputations required because no samples had 0 ppm lead

Appendix C: Comparing current eyeliner testing methods

Table C1: Linear regression results for *all eyeliners* (n=37).

Estimate	Adjusted R ²	Robust std error	95% CI	F stat	p-value	Pearson's correlation	Spearman's correlation
1.43	0.80	0.13	[1.16, 1.71]	113.52	< 0.001	0.89	0.86

Table C2: Linear regression results for *all eyeliners <400,000 ppm Pb* (n=23).

Estimate	Adjusted R ²	Robust std error	95% CI	F stat	p-value	Pearson's correlation	Spearman's correlation
2.295	0.94	0.21	[1.85, 2.74]	115.74	< 0.001	0.97	0.79

Table C3: Linear regression results for *powder eyeliners <400,000 ppm Pb* (n=10).

Estimate	Adjusted R ²	Robust std error	95% CI	F stat	p-value	Pearson's correlation	Spearman's correlation
2.34	0.93	0.22	[1.83, 2.86]	109.37	< 0.001	0.95	0.74

Table C4: Linear regression results for *all cream eyeliners <90 ppm Pb* (n=8).

Estimate	Adjusted R ²	Robust std error	95% CI	F stat	p-value	Pearson's correlation	Spearman's correlation
0.52	0.12	0.38	[-0.41, 1.44]	1.88	0.22	0.36	0.40

Table C5: Linear regression descriptive statistics for comparison between laboratory Pb measurements (ICP-MS, ICP-OES, GFAA) with average XRF Pb measurements for *eyeliners including <LOD measurements*.

Eyeliners analyzed	n	Method	% <RL or <LOD	Mean +/- SD (ppm)	s
<i>All eyeliners</i>	82	Laboratory	22% (18)	120,986 +/- 235,542	5.1 [<0.02, 785,500]
		XRF	54.5% (187)	193,005 +/- 376,674	10.8 [<0, 993,561]
<i>All eyeliners <400,000 ppm</i>	68	Laboratory	19% (13)	19,116 +/- 74,778	2.8 [<0.02, 391,500]
		XRF	55% (187)	47,060 +/- 178,866	8.4 [<0, 990,619]
<i>Powder eyeliners <400,000 ppm</i>	26	Laboratory	4% (1)	89,705 +/- 196,377	5.3 [<0.02, 391,500]
		XRF	33% (43)	162,893 +/- 170,259	9.98 [<0, 990,619]
<i>Cream eyeliners <90 ppm</i>	39	Laboratory	44% (17)	3.73 +/- 5.56	1.6 [<0.45, 29]
		XRF	67% (131)	9.78 +/- 10.54	9.98 [<0, 38.6]

Table C6: Linear regression results for *all eyeliners including <LOD measurements* (n=82).

Estimate	Adjusted R ²	Robust std error	95% CI	F stat	p-value	Pearson's correlation	Spearman's correlation
1.49	0.86	0.13	[1.23, 1.74]	132.9	< 0.001	0.93	0.64

Table C7: Linear regression results for *eyeliners <400,000 ppm Pb including <LOD measurements* (n=68).

Estimate	Adjusted R ²	Robust std error	95% CI	F stat	p-value	Pearson's correlation	Spearman's correlation
2.28	0.94	0.21	[1.86, 2.70]	117.36	< 0.001	0.97	0.38

Table C8: Linear regression results for *powder eyeliners <400,000 ppm Pb including <LOD measurements* (n=26).

Estimate	Adjusted R ²	Robust std error	95% CI	F stat	p-value	Pearson's correlation	Spearman's correlation
2.32	0.94	0.22	[1.84, 2.79]	114.19	< 0.001	0.97	0.74

Table C9: Linear regression results for *all cream eyeliners <90 ppm Pb including <LOD measurements* (n=39).

Estimate	Adjusted R ²	Robust std error	95% CI	F stat	p-value	Pearson's correlation	Spearman's correlation
0.81	0.16	0.33	[0.14, 1.48]	5.98	0.02	0.40	0.16

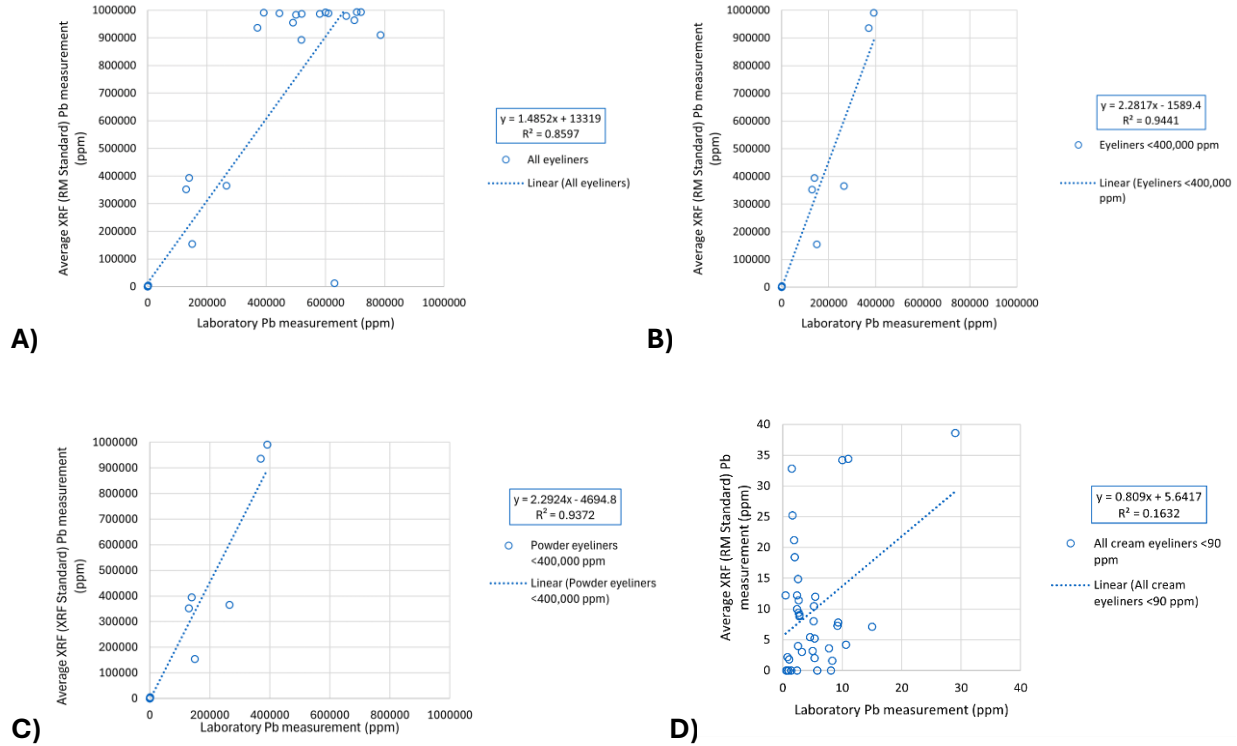


Figure C1: **A)** Linear regression to compare laboratory Pb measurements (ICP-MS, ICP-OES, GFAA) with average XRF (XRF RM Standard) Pb measurements *including <LOD measurements for A) all eyeliners (n=82). B) eyeliners <400,000 ppm (n=68). C) powdered eyeliners <400,000 ppm (n=26). D) cream eyeliners <90 ppm (n=39).*

Appendix D: XRF optimization

Table D1: Linear regression results for the three XRF methods for *all calibrants* (5 ppm – 842,600 ppm, n=11).

	Estimate	Adjusted R ²	Robust SE	95% CI	F stat	p-value	Pearson's correlation	Spearman's correlation
XRF RM Standard								
Calibrant	1.23	0.99	0.034	[1.17, 1.30]	1302.89	< 2.2e-16	0.99	0.99
XRF RM 180								
Calibrant	0.78	1.0	0.014	[0.75, 0.81]	3145.18	< 2.2e-16	1.0	1.0
XRF Soil								
Calibrant	1.12	0.99	0.033	[1.05, 1.19]	1108.78	< 2.2e-16	0.99	0.99

Table D2: Linear regression results for the three XRF methods for *calibrants <400,000* (5 ppm – 51,300 ppm, n=9).

	Estimate	Adjusted R ²	Robust SE	95% CI	F stat	p-value	Pearson's correlation	Spearman's correlation
<i>XRF RM Standard</i>								
Calibrant	0.53	1.0	0.034	[0.51, 0.56]	1912.37	< 2.2e-16	1.0	0.99
<i>XRF RM 180</i>								
Calibrant	0.42	1.0	1.901e-03	[0.42, 0.84]	49938.4	< 2.2e-16	1.0	0.99
<i>XRF Soil</i>								
Calibrant	0.87	1.0	0.01181	[0.82, 0.87]	5130.57	< 2.2e-16	1.0	0.99

Table D3: Linear regression results for all three XRF methods for *all eyeliners* (32 ppm – 720,000 ppm, n=12).

	Estimate	Adjusted R ²	Robust SE	95% CI	F stat	p-value	Pearson's correlation	Spearman's correlation
<i>XRF RM Standard</i>								
Lab	1.54	0.85	0.07	[1.32, 1.61]	419.31	< 2.2e-16	0.96	0.96
<i>XRF RM 180</i>								
Lab	1.38	0.60	0.1	[1.18, 1.57]	200.53	< 2.2e-16	0.78	0.87
<i>XRF Soil</i>								
Lab	1.62	0.86	0.06	[1.43, 1.66]	707.78	< 2.2e-16	0.98	0.93

Table D4: Linear regression results for all three XRF methods for *eyeliners <400,000 ppm Pb* (32 ppm – 391,500 ppm, n=6).

	Estimate	Adjusted R ²	Robust SE	95% CI	F stat	p-value	Pearson's correlation	Spearman's correlation
<i>XRF RM Standard</i>								
Lab	2.22	0.83	0.15	[1.93, 2.52]	234.27	< 0.001	0.91	0.94
<i>XRF RM 180</i>								
Lab	0.15	0.01	0.25	[-0.39, 0.68]	0.31	0.58	0.72	0.46
<i>XRF Soil</i>								
Lab	0.92	0.59	0.10	[0.71, 1.14]	77.92	< 0.001	0.77	0.80

Table D5: Linear regression results for the three XRF methods for *all calibrants including <LOD measurements* (5 ppm – 842,600 ppm, n=11).

	Estimate	Adjusted R ²	Robust SE	95% CI	F stat	p-value	Pearson's correlation	Spearman's correlation
<i>XRF RM Standard</i>								
Calibrant	1.23	0.99	0.034	[1.17, 1.30]	1314.71	< 2.2e-16	0.99	0.99
<i>XRF RM 180</i>								
Calibrant	0.78	1.0	0.014	[0.75, 0.81]	3177.58	< 2.2e-16	1.0	1.0
<i>XRF Soil</i>								
Calibrant	1.12	0.99	0.034	[1.05, 1.18]	1106.97	< 2.2e-16	0.99	0.99

Table D6: Linear regression results for all three XRF methods for *all eyeliners including <LOD measurements* (5.1 ppm – 720,000 ppm, n=13).

	Estimate	Adjusted R ²	Robust SE	95% CI	F stat	p-value	Pearson's correlation	Spearman's correlation
<i>XRF RM Standard</i>								
Lab	1.59	0.88	0.07	[1.45, 1.73]	514.54	< 2.2e-16	0.94	0.90
<i>XRF RM 180</i>								
Lab	1.42	0.66	0.08	[1.26, 1.58]	317.28	< 2.2e-16	0.81	0.90
<i>XRF Soil</i>								
Lab	1.62	0.88	0 0.0	[1.52, 1.72]	996.96	< 2.2e-16	0.94	0.94

Table D7: Linear regression results for all three XRF methods for *eyeliners <400,000 ppm including <LOD measurements* (5.1 ppm – 391,500 ppm, n=7).

	Estimate	Adjusted R ²	Robust SE	95% CI	F stat	p-value	Pearson's correlation	Spearman's correlation
<i>XRF RM Standard</i>								
Lab	2.24	0.85	0.14	[1.96, 2.52]	257.23	< 2.2e-16	0.92	0.96
<i>XRF RM 180</i>								
Lab	0.35	0.03	0.22	[-0.10, 0.80]	2.50	0.12	0.18	0.66
<i>XRF Soil</i>								

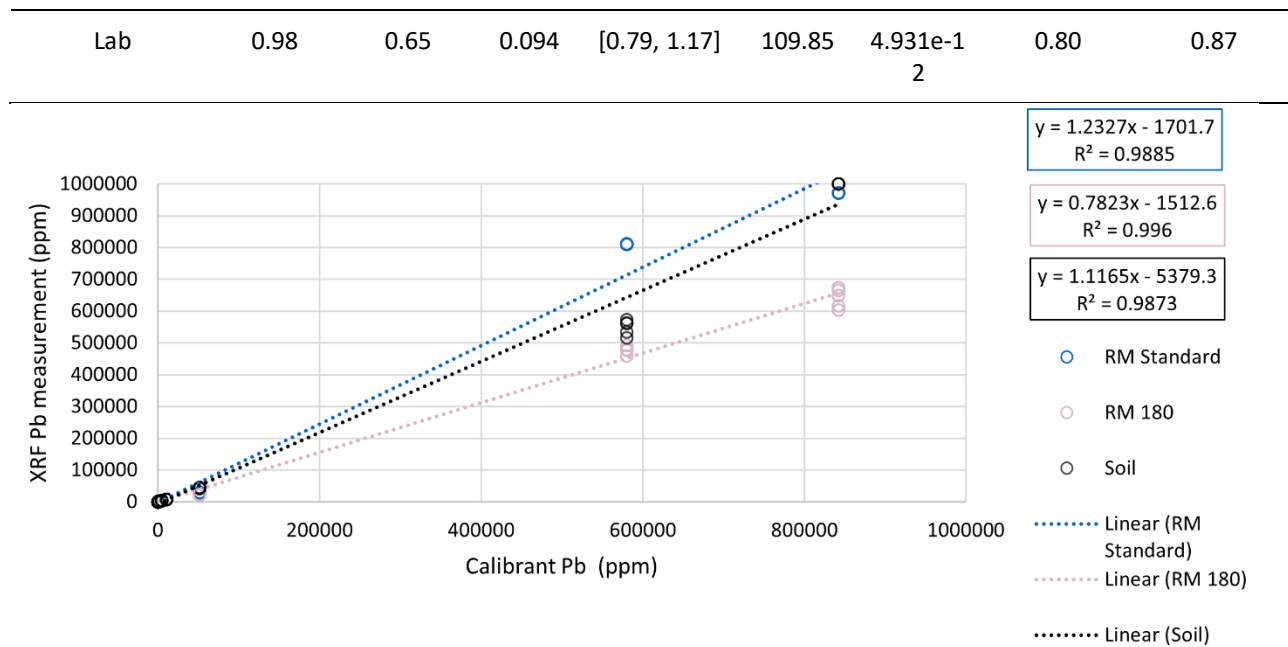


Figure D1: Linear regressions assessing the performance of the three XRF methods for *all calibrants including <LOD measurements* (5 ppm – 842,600 ppm, n=11).

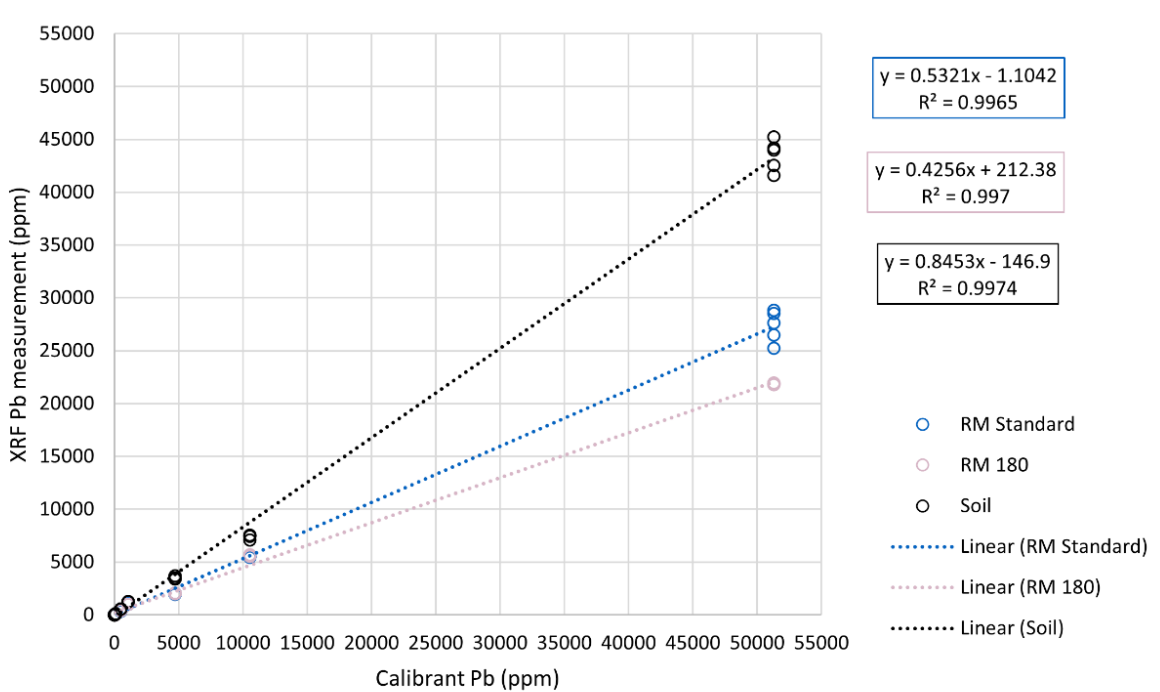


Figure D2: Linear regressions assessing the performance of the three XRF methods for *calibrants <400,000 ppm including <LOD measurements* (5 ppm – 51,300 ppm, n=9).

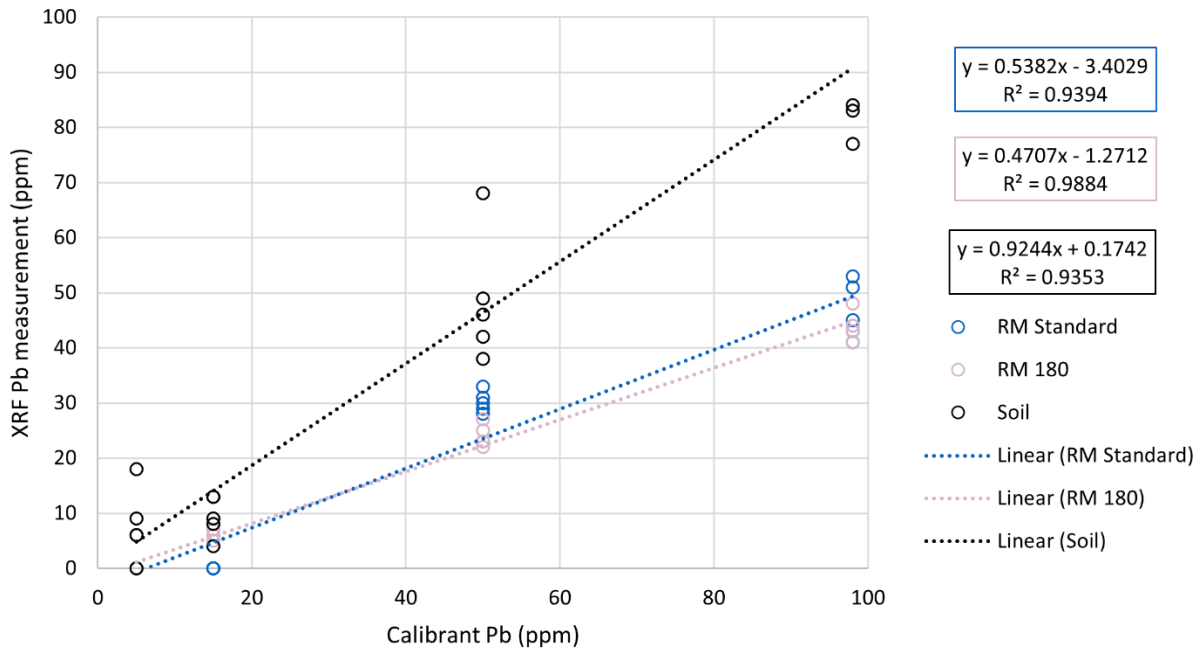


Figure D3: Linear regressions assessing the performance of the three XRF methods for **low lead calibrants including <LOD measurements** (5 ppm - 98 ppm, n=4). For more descriptive statistics see appendix.

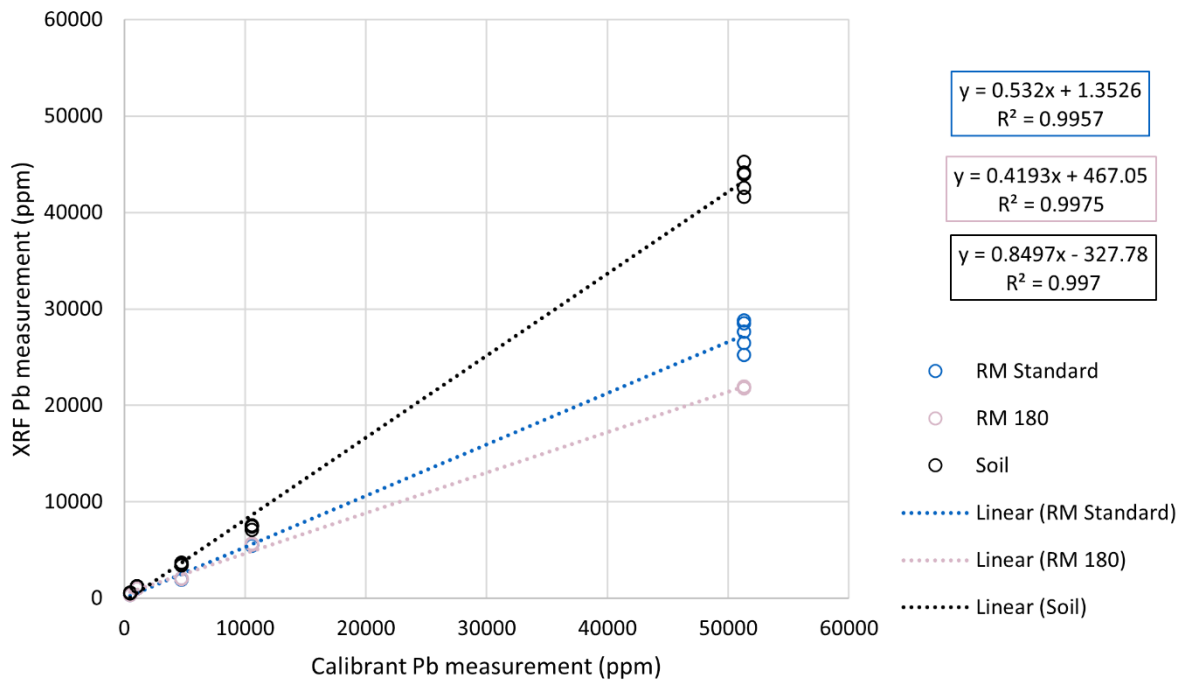


Figure D4: Linear regressions assessing the performance of the three XRF methods for **medium lead calibrants including <LOD measurements** (489 ppm - 51,300 ppm, n=5).

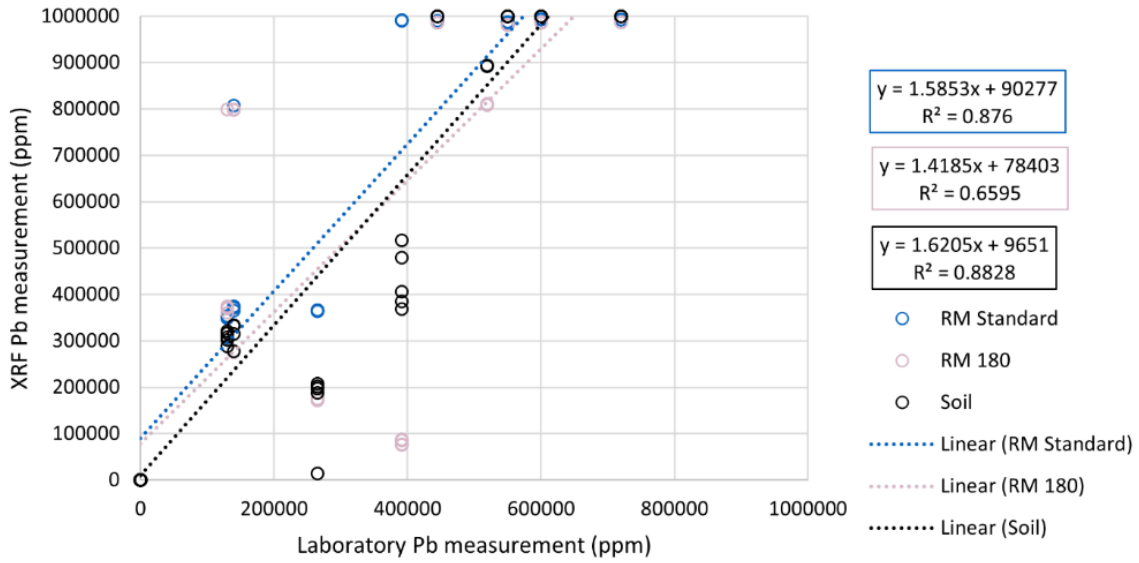


Figure D5: Linear regressions assessing the performance of the three XRF methods by comparing laboratory measurements with XRF methods for *all tested eyeliner including <LOD measurements* (GFAA (n=9), ICP-MS (n=4), 5.1 ppm - 720,000 ppm).

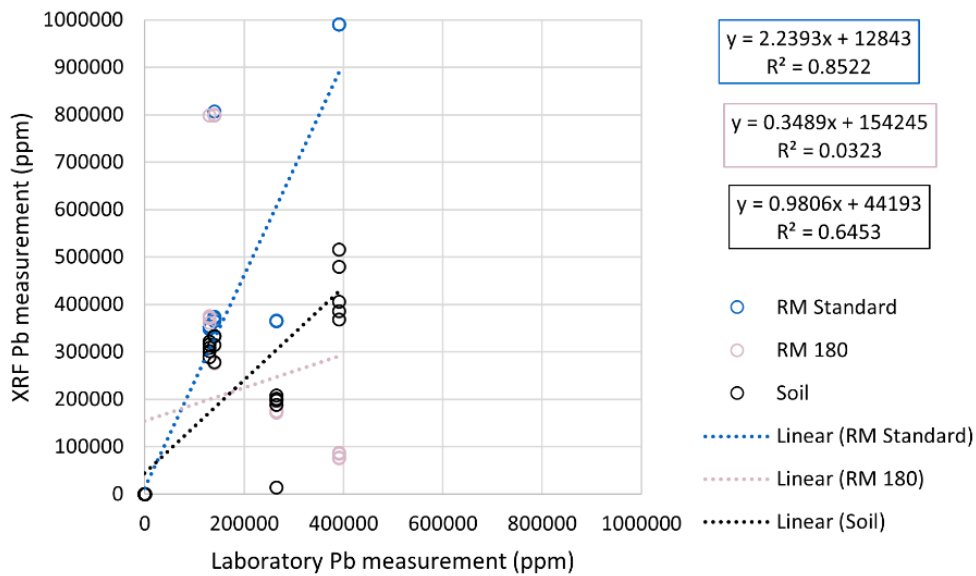


Figure D6: Linear regressions assessing the performance of the three XRF methods by comparing laboratory measurements with XRF methods for *all tested eyeliner <400,000 ppm including <LOD measurements* (GFAA (n=5), ICP-MS (n=2), 5.1 ppm - 391,500 ppm).

Appendix E: XRF Quality Checks

Table E1: XRF testing of standards provided by Bruker manufacturer during calibrant testing for Aim 3. All XRF Pb measurements fell within Bruker's acceptable Pb range.

Date of testing	Standard	Mean +/- SD (ppm)	Mean error +/- SD (ppm)
11/13/23	SAC305	783 +/- 202	213 +/- 7.3
11/13/23	SAC305	698 +/- 102	204 +/- 3.5
11/13/23	CS-M5	1,018 +/- 135	21 +/- 0.8
12/4/23	CS-M5	835 +/- 12.6	23.6 +/- 0.5

Table E2: XRF testing of standards provided by the Bruker manufacturer during eyeliner testing for Aim 3. All XRF Pb measurements fell within Bruker's acceptable Pb range.

Date of testing	Standard	Mean +/- SD (ppm)	Mean error +/- SD (ppm)
3/21/24	CS-M5	796 +/- 12.3	23 +/- 0
3/21/24	CS-M5	802 +/- 8.9	22.8 +/- 0.45

Bibliography

1. Tasslimi A.: Blood Lead Level (BLL) Screening: % Children with Elevated Blood Lead Levels, FFY 2016 to 2020, WA State. Personal communication between Azadeh Tasslimi (Washington State Department of Health) And Stephen G. Whittaker (Hazardous Waste Management Program in King County, Washington). Oct 2021.
2. "Blood Lead Levels in Children." Centers for Disease Control and Prevention. Retrieved from:
<https://www.cdc.gov/nceh/lead/prevention/blood-lead-levels.htm>.
3. "Screening for Lead during the Domestic Medical Examination for Newly Arrived Refugees ." Centers for Disease Control and Prevention. Retrieved from:
<https://www.cdc.gov/immigrantrefugeehealth/guidelines/lead-guidelines.html#:~:text=Top%20of%20Page-,Recommendations%20for%20Post%20Arrival%20Lead%20Screening,if%20currently%20pregnant%20or%20lactating.>
4. Ali M., A. Holland, S. Cohen, D. Vonnahme, M. Wilson, F. Amsalu, **et al**: *Elevated Blood Lead Levels in Refugee Children*
King County Public Health Partnership. , 2021. Retrieved from:
<https://www.kingcountyhazwastewa.gov/-/media/hazwaste/lhwmp-documents/technical-reports/rsh-king-county-public-health-partnership-refugee-report.pdf>.
5. Public Health- Seattle & King County (PHSKC): *2022 Product Testing Event*. , 2023. Retrieved from:
<https://kingcounty.gov/depts/health/environmental-health/toxins-air-quality/arsenic->

[lead/~media/depts/health/environmental-health/documents/toxins/product-testing-results-report.ashx.](#)

6. "Limiting Lead in Lipstick and Other Cosmetics." US Food and Drug Administration. Retrieved from:

[https://www.fda.gov/cosmetics/cosmetic-products/limiting-lead-lipstick-and-other-cosmetics.](https://www.fda.gov/cosmetics/cosmetic-products/limiting-lead-lipstick-and-other-cosmetics)

7. Mena S.: Toxic-Free Cosmetics Act (RCW 70A.560.020). (2023) (2023).

8. "CDC's Recommended Terminology When Discussing Children's Blood Lead Levels." Centers for Disease Control and Prevention. Retrieved from: [https://www.cdc.gov/nceh/lead/docs/CDCs-](https://www.cdc.gov/nceh/lead/docs/CDCs-recommended-terminology-discussing-childrens-BLLs-508.pdf)

[recommended-terminology-discussing-childrens-BLLs-508.pdf.](https://www.cdc.gov/nceh/lead/docs/CDCs-recommended-terminology-discussing-childrens-BLLs-508.pdf)

9. Mahmood Z.A.: Kohl Use in Antiquity Effects on the Eye. (2015).

10. Hankir Z.: *Eyeliners: A Cultural History*. : Vintage Digital, 2023.

11. At-Tirmidhi, Abi 'Eesaa Muhammad bin 'Eesah bin Sorah: Chapter on the kuhl of Sayyidina Rasulullah Sallallahu 'Alayhi Wasallam. In *Shamaa-Il Tirmidhi*, Anonymous , pp. 51-55: Inter-Islam, 1998.

12. Pedram R. and N. Pedram: Routine study coordination meeting. (2023).

13. Pirzada S., Z. Anwar, N. Iqbal, R. Gul, F. Noor, Z. Anwar, **et al**: Eye Care Practices and use of Surma in Newborns.

14. Anwar Z., S. Anwar, M. Akram, S. Ahmad, S. Pirzada, and A. Kanwal: Eye Care and Surma Use in Pakistani Neonates. *Pakistan Journal of Medical and Health Sciences* 16(2):374-375 (2022).

15. "Toxicological Profile for Lead." Agency for Toxic Substances and Disease Registry. Retrieved from:

[https://wwwn.cdc.gov/TSP/ToxProfiles/ToxProfiles.aspx?id=96&tid=22.](https://wwwn.cdc.gov/TSP/ToxProfiles/ToxProfiles.aspx?id=96&tid=22)

16. Sadeq M., T. Benamar, and A. Facciorusso: Is the application of Kohl to eyes associated with increased blood lead levels in children? A meta-analysis. *PAMJ - One Health* 4 (2021).
17. Navarro-Tapia E., M. Serra-Delgado, L. Fernández-López, M. Meseguer-Gilabert, M. Falcón, G. Sebastiani, **et al**: Toxic Elements in Traditional Kohl-Based Eye Cosmetics in Spanish and German Markets. *International journal of environmental research and public health* 18(11):6109 (2021).
18. "Kajal." utsavpedia. Retrieved from: <https://www.utsavpedia.com/cultural-connections/defining-traditional-vision-kohl/>.
19. "How to make kajal (surma) at home: kan-mai preparation." Ramya Venkateshwaran. Retrieved from: https://www.wildturmeric.net/homemade-kohlkajal/#google_vignette.
20. United States Food and Drug Administration, (FDA): "Detention Without Physical Examination Of Cosmetics That are Adulterated and/or Misbranded Due to Color Additive Violations". (2024).
21. Filella M., A. Martignier, and A. Turner: Kohl containing lead (and other toxic elements) is widely available in Europe. *Environmental research* 187:109658 (2020).
22. Hardy A.D., H.H. Sutherland, and R. Vaishnav: A study of the composition of some eye cosmetics (kohls) used in the United Arab Emirates. *Journal of Ethnopharmacology*80(2-3):137 (2002).
23. Daar E., A. Alsubaie, K.S. Almugren, S. Barnes, A. Alanazi, A. Alyahyawi, **et al**: X-ray fluorescence analysis of Pb, Fe and Zn in kohl. *Results in Physics* 7:3590-3595 (2017).
24. "Kohl (cosmetics)." Wikipedia. Retrieved from: [https://en.wikipedia.org/wiki/Kohl_\(cosmetics\)](https://en.wikipedia.org/wiki/Kohl_(cosmetics)).

25. "Surma: Can't you see how bad it is?" Samra Aamir, Myra Iqbal and Naveela Khan. Retrieved from: <https://tribune.com.pk/story/599470/surma-cant-you-see-how-bad-it-is>.
26. O. Rasheed B., S. M. Ali, and P. M. Aljaff: Microbial and Heavy Metal Contamination in Five Brands of Kohl Used in Sulaimani City. *Journal of Zankoy Sulaimani - Part A* 23(1):94-100 (2021).
27. Buksh E., S.A. Naz, A. Zubair, K. Yasmeen, M. Shafique, N. Jabeen, **et al**: Kohl: A Widely used eye Cosmetic with Hazardous Biochemical Composition. *Biosciences, biotechnology research Asia* 17(3):621-628 (2020).
28. Mahmood Z.A., A. Saeed, K. Usmanghani, O. Ali, S. Jahan, R. Zaihd, **et al**: Kohl (surma): Retrospect and prospect. *Pakistan journal of pharmaceutical sciences* 22(1):107-122 (2009).
29. "What are Oriental Surma Kohl Eyeliners Made Out of?" Raqib Zaman. Retrieved from: <https://www.therevisionist.org/reviews/what-are-oriental-surma-kohl-eyeliners-made-out-of/#:~:text=Rather%2C%20it%20is%20commonly%20Zincite,UV%20rays%20of%20the%20sun>.
30. Mohta A.: Kajal (Kohl) - A dangerous cosmetic. *Oman Journal of Ophthalmology* 3(2):100-101 (2010).
31. Khan, Ullah, Khan, Zafar, Farhan Ali Khan, Mustaqeem, **et al**: Personal-Care Cosmetic Practices in Pakistan: Current Perspectives and Management. *CCID* Volume14:9 (2021).
32. Al-Ashban R.M., M. Aslam, and A.H. Shah: Kohl (surma): a toxic traditional eye cosmetic study in Saudi Arabia. *Public Health* 118(4):292-298 (2004).
33. McMichael J., J. McMichael, B. Stoff, and B. Stoff: Surma eye cosmetic in Afghanistan: a potential source of lead toxicity in children. *Eur J Pediatr* 177(2):265-268 (2018).

34. "Sunan and Manners of Applying Kohl." Dawat-e-Islami. Retrieved from:
<https://www.dawateislami.net/blog/sunan-and-manners-of-applying-kohl>.
35. Ali A.R., O.R. Smales, and M. Aslam: Surma and lead poisoning. *BMJ* 2(6142):915-916 (1978).
36. Sweha F.: Kohl along history in medicine and cosmetics. *Histoire des Sciences médicales* 17:182-183 (1982).
37. Tiffany-Castiglioni E., R. Barhoumi, and Y. Mouneimne: Kohl and surma eye cosmetics as significant sources of lead (Pb) exposure. *Journal of Local and Global Health Science* 2012(2012) (2012).
38. Patel A. and A. Athawale: Blood lead levels in children with encephalopathy. *Indian pediatrics* 46(10):845-848 (2009).
39. Sprinkle R.V.: Leaded eye cosmetics: a cultural cause of elevated lead levels in children. *The Journal of family practice* 40(4):358-362 (1995).
40. RAHBAR M.H., F. WHITE, M. AGBOATWALLA, S. HOZHABRI, and S. LUBY: Factors associated with elevated blood lead concentrations in children in Karachi, Pakistan. *Bulletin of the World Health Organization* 80(10):769-775 (2002).
41. Begay L Crystal: Childhood Lead Exposure Associated with the Use of Kajal, an Eye Cosmetic from Afghanistan — Albuquerque, New Mexico, 2013
. *CDC Morbidity and Mortality Report* 14(1):e23-e24 (2013).
42. Morbidity and Mortality Weekly Report (MMWR). (2012).

43. Lamas G.A., A. Bhatnagar, M.R. Jones, K.K. Mann, K. Nasir, M. Tellez-Plaza, **et al**: Contaminant Metals as Cardiovascular Risk Factors: A Scientific Statement From the American Heart Association. *Journal of the American Heart Association* 12:e029852 (2023).
44. The UN Refugee Agency: *UNHCR's Executive Committee Conclusion on Children at Risk*. : The UN Refugee Agency, 2007. Retrieved from: <https://www.unhcr.org/publications/unhcrs-executive-committee-conclusion-children-risk>.
45. Balza, Bikomeye, Beyer, Rublee, and Flynn: Elevated blood lead levels of refugee children in the United States: a systematic review of recent literature (2011–2021). *Reviews on environmental health* 38(2):361-383 (2023).
46. Chatham-Stephens, Caravanos, Ericson, Landrigan, and Fuller: The pediatric burden of disease from lead exposure at toxic waste sites in low and middle income countries. *Environmental research* 132:379-383 (2014).
47. Fewtrell, Prüss-Üstün, Landrigan, and Ayuso-Mateos: Estimating the global burden of disease of mild mental retardation and cardiovascular diseases from environmental lead exposure. *Environmental Research* 94(2):120 (2004).
48. Uniform Washington Food, Drug, and Cosmetic Act. (1945).
49. Mean S., C. Ryu, L. Berry, T. Simmons, D. Duerr, R. Goodman, **et al**: Concerning the use of toxic chemicals in cosmetic products. (2023) (2023).
50. Regulation (EC) No 1223/2009 Of the European Parliament and of the Council of 30 November 2009 on cosmetic products. 14 (2009).

51. Bund B.: Technically avoidable heavy metal contents in cosmetic products. *J Consum Prot Food Saf* 12(1):51-53 (2017).
52. Government of Canada: Guidance on Heavy Metal Impurities in Cosmetics. (2012).
53. Ministry of Health and Family Welfare: The Drugs and Cosmetics Act and Rules. (1940).
54. Gulf Technical Regulation: Cosmetic Products – Safety Requirements of Cosmetics and Personal Care Products. (2021).
55. Personal communication between Ali Mohamed, Aesha Mokashi, and Stephen G Whittaker during AHI routine study coordination meeting. (2023).
56. Gouitaa H., A. Bellaouchou, M. Fekhaoui, A.E. Abidi, N. Mahnine, and R.B. Aakame: Assessment of lead levels in traditional eye cosmetic "kohl" frequently used in Morocco and Health hazard. 7(2) (2016).
57. "Metal Content of Consumer Products Tested by the NYC Health Department." "Metal Content of Consumer Products Tested by the NYC Health Department." Retrieved from: <https://data.cityofnewyork.us/Health/Metal-Content-of-Consumer-Products-Tested-by-the-N/da9u-wz3r>.
58. "Lead Content of Consumer Products tested in King County, Washington ." Public Health – Seattle & King County and the Hazardous Waste Management Program. Retrieved from: https://data.kingcounty.gov/Health-Wellness/Lead-Content-of-Consumer-Products-tested-in-King-C/i6sy-ckp7/about_data.
59. Charter E., M. Duekman, A. Kurychack, B. Martyn, N. Nnebe, and B. Raymer: The Health Risks of Hidden Heavy Metals in Face Makeup . (2011).

60. Sainio E., R. Jolanki, E. Hakala, and L. Kanerva: Metals and arsenic in eye shadows. 42(1):5-10 (2000).

61. Pawlaczyk A., M. Gajek, M. Balcerek, and M.I. Szykowska-Jóźwik: Determination of Metallic Impurities by ICP-MS Technique in Eyeshadows Purchased in Poland. Part I. 26(21):6753 (2021).

62. Volpe M.G., M. Nazzaro, R. Coppola, F. Rapuano, and R.P. Aquino: Determination and assessments of selected heavy metals in eye shadow cosmetics from China, Italy, and USA. 101:65-69 (2012).

63. Al-Saleh I., S. Al-Enazi, and N. Shinwari: Assessment of lead in cosmetic products. 54(2):105-113 (2009).

64. "Comparison of ICP-OES and ICP-MS for Trace Element Analysis

." "Comparison of ICP-OES and ICP-MS for Trace Element Analysis

." Retrieved from:

<https://www.thermofisher.com/us/en/home/industrial/environmental/environmental-learning-center/contaminant-analysis-information/metal-analysis/comparison-icp-oes-icp-ms-trace-element-analysis.html>.

65. "Afghan Health Initiative." "Afghan Health Initiative." Retrieved from: <https://afghanhealth.org/our-mission>.

66. "Is My Handheld XRF Analyzer Working Correctly?" Bruker. Retrieved from:

<https://www.bruker.com/en/products-and-solutions/elemental-analyzers/handheld-xrf-spectrometers/is-my-handheld-xrf-analyzer-working.html>.

67. U.S. Environmental Protection Agency: Method 6020A Inductively Coupled Plasma-Mass Spectrometry. Revised 2007. (2007).

68. United States Consumer Product Safety Commission Directorate for Laboratory Sciences Division of Chemistry: United States Consumer Product Safety Commission Test Method: CPSC-CH-E1002-08.3 Standard Operating Procedure for Determining Total Lead (Pb) in Nonmetal Children's Products, Revision . (2012).

69. United States Environmental Protection Agency (: Method 7010 Graphite Furnace Atomic Absorption Spectrophotometry. (1998).

70. Ceballos D.M., R.F. Herrick, Z. Dong, A. Kalweit, M. Miller, J. Quinn, **et al**: Factors affecting lead dust in construction workers' homes in the Greater Boston Area. *Environmental Research*195 (2020).

71. "Home CRAN NADA2 Censored Data Analysis" Censored Data Analysis". "Home CRAN NADA2 Censored Data Analysis" Censored Data Analysis". Retrieved from:
https://rdrr.io/cran/NADA2/f/vignettes/Censored_Data_Analysis.Rmd.

72. United States Environmental Protection Agency (: SW-846 Test Method 6200: Field Portable X-Ray Fluorescence Spectrometry for the Determination of Elemental Concentrations in Soil and Sediment. (2007).

73. Omolaoye T.S. and S.S. du Plessis: Heavy metal assessment of some eye shadow products imported into Nigeria from China 93 (2020).

74. Lim D.S., T.H. Roh, M.K. Kim, Y.C. Kwon, S.M. Choi, S.J. Kwack, **et al**: Non-cancer, cancer, and dermal sensitization risk assessment of heavy metals in cosmetics. 81(11):432-452 (2018).

75. Omenka S.S. and A.A. Adeyi: Heavy metal content of selected personal care products (PCPs) available in Ibadan, Nigeria and their toxic effects. 3:628-635 (2016).