VISUAL ANALYTICS OF ARSENIC IN VARIOUS FOODS

by

MATILDA OLUBUNMI JOHNSON

submitted in accordance with the requirements
for the degree of

DOCTOR OF PHILOSOPHY

in the subject

ENVIRONMENTAL SCIENCE

at the

UNIVERSITY OF SOUTH AFRICA

SUPERVISOR: DR C YEDJOU

CO- SUPERVISORS: PROF O R AWOFOLU, PROF S J MOJA

JUNE 2014

VISUAL ANALYTICS OF ARSENIC IN VARIOUS FOODS

By

MATILDA OLUBUNMI JOHNSON

MS, Texas Woman's University, Denton TX, USA
BSc, University of Ibadan, Nigeria

DECLARATION

I declare that “Visual Analytics of Arsenic in Various Foods” is my own work, and that all the sources that I have used or quoted have been indicated and acknowledged by means of complete references.

Matilda O Johnson

Signature and date: June 9, 2014

Student

DEDICATION

I dedicate this thesis to my husband and my children for their love and support through all this. I also dedicate it to my parents and my siblings who have always supported me with their prayers, love, and encouragement.

ABSTRACT

Arsenic is a naturally occurring toxic metal and its presence in food composites could be a potential risk to the health of both humans and animals. Arsenic-contaminated groundwater is often used for food and animal consumption, irrigation of soils, which could potentially lead to arsenic entering the human food chain. Its side effects include multiple organ damage, cancers, heart disease, diabetes mellitus, hypertension, lung disease and peripheral vascular disease. Research investigations, epidemiologic surveys and total diet studies (market baskets) provide datasets, information and knowledge on arsenic content in foods. The determination of the concentration of arsenic in rice varieties is an active area of research. With the increasing capability to measure the concentration of arsenic in foods, there are volumes of varied and continuously generated datasets on arsenic in food groups. Visual analytics, which integrates techniques from information visualization and computational data analysis via interactive visual interfaces, presents an approach to enable data on arsenic concentrations to be visually represented.

The goal of this doctoral research in Environmental Science is to address the need to provide visual analytical decision support tools on arsenic content in various foods with special emphasis on rice. The hypothesis of this doctoral thesis research is that software enabled visual representation and user interaction facilitated by visual interfaces will help discover hidden relationships between arsenic content and food categories.

The specific objectives investigated were: (1) Provide insightful visual analytic views of compiled data on arsenic in food categories; (2) Categorize table ready foods by arsenic content; (3) Compare arsenic content in rice product categories and (4) Identify informative sentences on arsenic concentrations in rice. The overall research

method is secondary data analyses using visual analytics techniques implemented through Tableau Software.

Several datasets were utilized to conduct visual analytical representations of data on arsenic concentrations in foods. These consisted of (i) arsenic concentrations in 459 crop samples; (ii) arsenic concentrations in 328 table ready foods from multi-year total diet studies; (iii) estimates of daily inorganic arsenic intake for 49 food groups from multi-country total diet studies; (iv) arsenic content in rice product categories for 193 samples of rice and rice products; (v) 758 sentences extracted from PubMed abstracts on arsenic in rice.

Several key insights were made in this doctoral research. The concentration of inorganic arsenic in instant rice was lower than those of other rice types. The concentration of Dimethylarsinic Acid (DMA) in wild rice, an aquatic grass, was notably lower than rice varieties (e.g. 0.0099 ppm versus 0.182 for a long grain white rice). The categorization of 328 table ready foods into 12 categories enhances the communication on arsenic concentrations. Outlier concentration of arsenic in rice were observed in views constructed for integrating data from four total diet studies. The 193 rice samples were grouped into two groups using a cut-off level of 3 mcg of inorganic arsenic per serving. The visual analytics views constructed allow users to specify cut-off levels desired. A total of 86 sentences from 53 PubMed abstracts were identified as informative for arsenic concentrations. The sentences enabled literature curation for arsenic concentration and additional supporting information such as location of the research. An informative sentence provided global “normal” range of 0.08 to 0.20 mg/kg for arsenic in rice. A visual analytics resource developed was a dashboard that facilitates the interaction with text and a connection to the knowledge base of the PubMed literature database.

The research reported provides a foundation for additional investigations on visual analytics of data on arsenic concentrations in foods. Considering the massive and complex data associated with contaminants in foods, the development of visual analytics tools are needed to facilitate diverse human cognitive tasks. Visual analytics tools can provide integrated automated analysis; interaction with data; and data visualization critically needed to enhance decision making. Stakeholders that would benefit include consumers; food and health safety personnel; farmers; and food producers. Arsenic content of baby foods warrants attention because of the early life exposures that could have life time adverse health consequences.

The action of microorganisms in the soil is associated with availability of arsenic species for uptake by plants. Genomic data on microbial communities presents wealth of data to identify mitigation strategies for arsenic uptake by plants. Arsenic metabolism pathways encoded in microbial genomes warrants further research. Visual analytics tasks could facilitate the discovery of biological processes for mitigating arsenic uptake from soil.

The increasing availability of central resources on data from total diet studies and research investigations presents a need for personnel with diverse levels of skills in data management and analysis. Training workshops and courses on the foundations and applications of visual analytics can contribute to global workforce development in food safety and environmental health. Research investigations could determine learning gains accomplished through hardware and software for visual analytics. Finally, there is need to develop and evaluate informatics tools that have visual analytics capabilities in the domain of contaminants in foods.

Keywords: arsenic, dietary, cancer, foods, rice, text mining, visual analytics

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

ABBREVIATIONS AND ACRONYMS

µg	Microgram
As	Inorganic arsenic
As3MT	Arsenic methyltransferase
AsIII	Arsenite
AsV	Arsenate
CCA	Chromated copper arsenate
CTD	Comparative Toxicogenomics Database
DMA	Dimethylarsenic Acid
DMAsIII	Dimethylarsonous acid
DMAsV	Dimethylarsonic acid
dw	Dry weight
ESFA	European Food Safety Authority
FAO	Food and Agriculture Organization
FDA	Food and Drug Administration
FSAI	Food Safety Authority of Ireland
JECFA	Joint Expert Committee on Food Additives
MAsIII	Methylarsonous acid
MAsV	Methylarsonic acid
MeSH	Medical Subject Heading
mcg	Microgram
MMA(III)	Monomethylarsonous acid
MMA(V)	Monomethylarsonic acid
PMID	PubMed Identifier
ppb	Parts per billion
PTWI	Permissible Human Weekly Exposure

PubMed	PubMed Literature Database
SAM	S-adenosylmethionine
TDS	Total Diet Study
TMAs(V)O	Trimethylarsine oxide
TMAsIII	Trimethylarsine
USEPA	United States Environmental Protection Agency
WHO	World Health Organization
Ww	Wet weight

ACKNOWLEDGEMENTS

First of all, I acknowledge and give thanks to Almighty God without whom I could never have achieved this doctoral thesis. I am very grateful to all who have helped me throughout my research and completion of this thesis research. I particularly thank my advisor Dr. Clement Yedjou.

I would also like to thank my co-advisor Prof. O. Awofolu for his guidance, encouragement and support. I am grateful to Prof. S. Moja and staff of the Department of Environmental Science, University of South Africa.

I thank the reviewers and co-authors of the first publication that resulted from this thesis titled "The case for visual analytics of arsenic concentrations in foods" and published in the International Journal of Environmental Research and Public Health, 7(5), 1970-1983, 2010.

CHAPTER 1

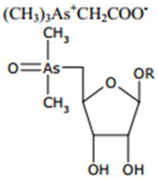
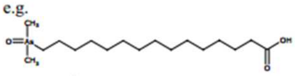
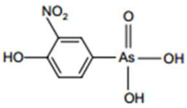
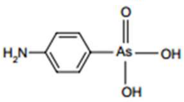
INTRODUCTION

1.1 Overview

Arsenic is a naturally occurring toxic metal and its presence in food composites could be a potential risk to the health of both humans and animals (Al Rmalli et al., 2005, Rintala et al., 2014, Zhao et al., 2010). Arsenic, a derivative the Greek word 'arsenikon' meaning 'potent' has been shown to be a very toxic element, particularly inorganic arsenic (Jolliffe, 1993). Arsenic toxicity affects millions of people worldwide and is a human carcinogen (Faita et al., 2013, Otles and Cagindi, 2010). Inorganic arsenic occurs naturally in soil, air and water as well as through anthropogenic sources such as mining, agriculture and non-agricultural activities (Duker et al., 2005, Garelick et al., 2008). The arsenic species and their relevance in foods are presented in Figure 1 obtained from a report of European Food Safety Authority on arsenic in food (European Food Safety Authority, 2009).

Arsenic-contaminated groundwater is often used for food and animal consumption, irrigation of soils, which could potentially lead to arsenic entering the human food chain (Al Rmalli et al., 2005). Its side effects include multiple organ damage, cancers, heart disease, diabetes mellitus, hypertension, lung disease and peripheral vascular disease (Faita et al., 2013).

Visual Analytics is a multidisciplinary field that is defined as the science of analytical reasoning facilitated by interactive visual interfaces (Keim et al., 2008, Keim et al., 2010, Thomas and Cook, 2005, Thomas and Cook, 2006). Furthermore, visual analytics combines techniques from *computer-based* *information visualization* with techniques from *computational transformation and analysis of data*. Research areas related to visual analytics are summarized in Figure 2.

Name	Abbreviation	Chemical structure[(a)]	Relevance/comment
Inorganic arsenic	iAs		Sum of As(III) and As(V).
Arsenite	As(III)	$As(O^-)_3$	Trace to low levels in most foods; highly toxic.
Arsenate	As(V)	$O{=}As(O^-)_3$	Trace to low levels in most foods; a major form in water; highly toxic.
Arsenobetaine	AB	$(CH_3)_3As^+CH_2COO^-$	Major arsenic species in most seafoods; non-toxic.
Arsenosugars[(b)]			Major (edible algae) or significant (molluscs) arsenic species in many seafoods.
Arsenolipids[(c)]		e.g. 	Newly discovered arsenic species present in fish oils and fatty fish; likely to be present in other seafoods as well.
Trimethylarsonio propionate	TMAP	$(CH_3)_3As^+CH_2CH_2COO^-$	Minor arsenic species present in most seafoods.
Methylarsonate	MA	$CH_3AsO(O^-)_2$	Trace arsenic species of some seafoods and terrestrial foods; a significant human urine metabolite of iAs.
Methylarsonite	MA(III)	$CH_3As(O^-)_2$	Not usually detected in foods; detected in some human urine samples as a metabolite of iAs; a toxic species thought to be important for arsenic's mode of toxic action.
Dimethylarsinate	DMA	$(CH_3)_2AsO(O^-)$	Minor arsenic species in seafoods and some terrestrial foods; the major human urine metabolite of iAs, arsenosugars and arsenolipids.
Thio-dimethylarsinate	Thio-DMA	$(CH_3)_2AsS(O^-)$	A minor human urine metabolite of inorganic arsenic and arsenosugars.
Dimethylarsinite	DMA(III)	$(CH_3)_2AsO^-$	Not detected in foods; detected in some human urine samples as a metabolite of iAs; a very unstable (reactive) species that is very difficult to measure; highly toxic species considered by some researchers to be central to arsenic's mode of toxic action.
Trimethylarsine oxide	TMAO	$(CH_3)_3AsO$	Minor arsenic species common in seafood.
Tetramethylarsonium ion	TETRA	$(CH_3)_4As^+$	Minor arsenic species common in seafood.
Arsenocholine	AC	$(CH_3)_3As^+CH_2CH_2OH$	Trace arsenic species found in seafood; is readily oxidised to arsenobetaine in biological systems.
Roxarsone			Used in the United States of America as a poultry feed additive to enhance growth; banned in Europe; not usually detected in food.
Arsanilic acid			Previously used as a drug and as an animal food additive; also used as its sodium salt (atoxyl).

(a): The simpler arsenic species are also often referred to in their protonated forms such as As(III) arsenous acid, H_3AsO_3; As(V) arsenic acid, H_3AsO_4; MA methylarsonic acid, $CH_3AsO(OH)_2$; DMA dimethylarsinic acid $(CH_3)_2AsO(OH)$; MA(III) methylarsonous acid $CH_3As(OH)_2$; DMA(III) dimethylarsinous acid $(CH_3)_2AsOH$.

(b): Over 20 arsenosugars have been reported as natural products; they differ by having different R groups on the aglycone portion of the molecule, and by replacing the oxygen on the arsenic atom with either a sulfur atom or a third methyl group (see Francesconi and Edmonds (1997)). Most of the arsenic present as arsenosugars, however, is contained in just four compounds based on the structure drawn above and with (i) $R{=}CH_2CHOHCH_2OH$; (ii) $R{=}CH_2CHOHCH_2OP(O)(OH)OCH_2CHOHCH_2OH$; (iii) $R{=}CH_2CHOHCH_2OSO_3H$; and (iv) $R{=}CH_2CHOHCH_2SO_3H$

(c): Nine arsenolipids have been reported so far (2009) as natural products, all of which contain the dimethylarsinoyl group $[(CH_3)_2As(0)\text{-}]$ bound to either one of several long chain fatty acids, or to long chain hydrocarbons. Many more arsenolipids are present in foods – their structures are currently unknown.

Figure 1. Names, abbreviations and chemical structure of arsenic species and their relevance in foods.

Source: (European Food Safety Authority, 2009)

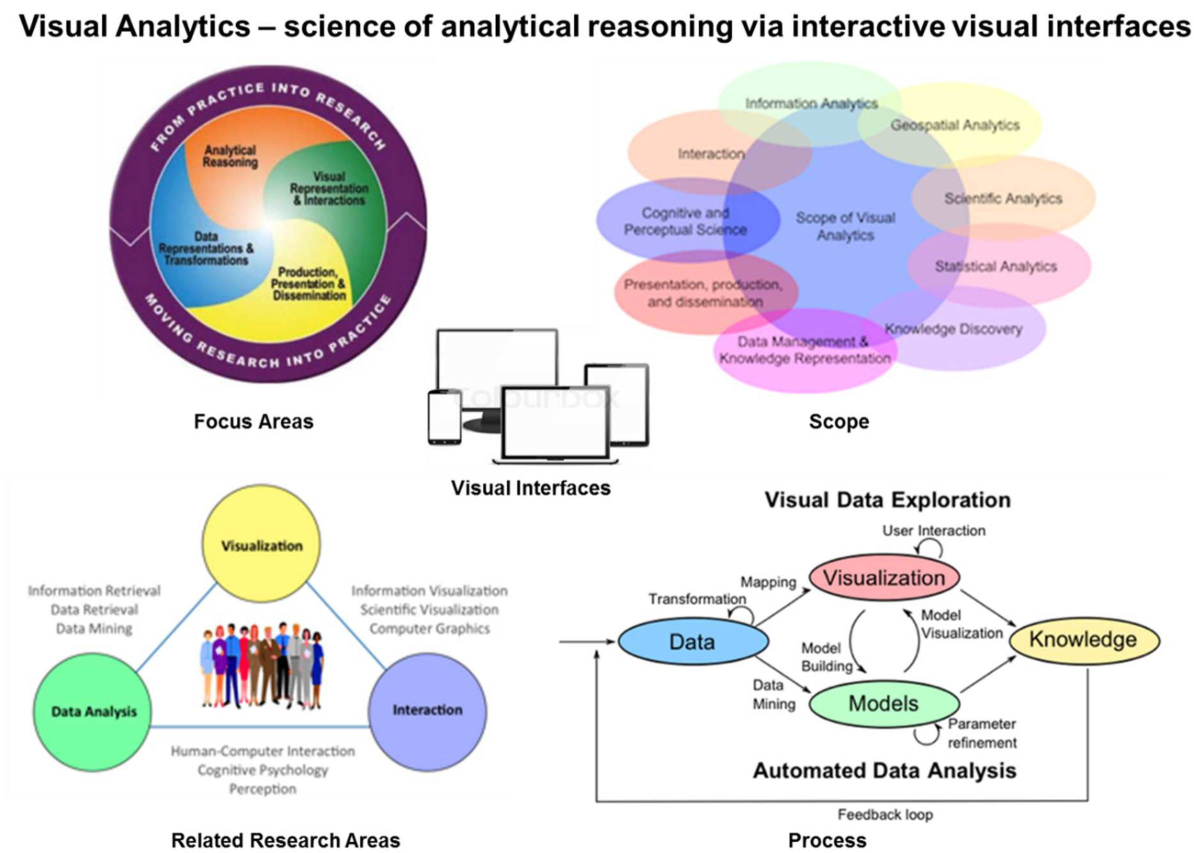


Figure 2. Visual Analytics - focus, scope, visual interfaces, related research areas and process

Notes:
Definition: The science of analytical reasoning via interactive visual interfaces.

Focus Areas: Analytical Reasoning; Visual Representations and Interactions; Data Representations and Transformations; Techniques for Production, Presentation and Dissemination.

Scope: Information Analytics; Geospatial Analytics; Scientific Analytics; Statistical Analytics; Knowledge Discovery; Data Management & Knowledge Representation; Presentation, Production and Dissemination; Cognitive and Perceptual Science; and Interaction.

Related Areas: Visual Analytics can be seen as an integral approach combining visualization, human factors, and data analysis. Besides visualization and data analysis, especially human factors, including the areas of cognition and perception, play an important role in the communication between the human and the computer, as well as in the decision-making process. With respect to visualization, Visual Analytics relates to the areas of Information Visualization and Computer Graphics, and with respect to data analysis, it profits from methodologies developed in the fields of information retrieval, data management & knowledge representation as well as data mining.

Process: The Visual Analytics Process combines automatic and visual analysis methods with a tight coupling through human interaction in order to gain knowledge from data. The figure shows an abstract overview of the different stages (represented through ovals) and their transitions (arrows) in the Visual Analytics Process.
Sources: (Keim et al., 2008, Keim et al., 2010, Thomas and Cook, 2005)

1.2. Motivation for Research

1.2.1 Regulatory Limits for Arsenic in Foods

Arsenic, particularly inorganic arsenic has been recognized as a toxic element, a carcinogen with a global impact on the health of humans and its intake must be limited (Naujokas et al., 2013, Tchounwou et al., 2003). The concentrations of arsenic vary in different foods; also it is not always possible to distinguish the form of arsenic in a food. This makes it impractical, almost impossible to provide regulatory limit for each food.

Several regulatory bodies worldwide including the Joint FAO/WHO Expert Committee on Food Additive, Food Standards Australia New Zealand, World Health Organization and United States Environmental Protection Agency have set various guideline levels for total arsenic, inorganic arsenic and organic arsenic levels in various foods and drinking water (Hite, 2013, JEFCA, 1989, JEFCA, 2011, JEFCA, 1995). These regulatory levels are expected to help consumers, risk managers, policy makers and responsible authorities minimize exposure of humans and animals to this toxic element.

1.2.2 Enhancing Decision-Making of Consumers on Arsenic in Foods

Research investigations, epidemiologic surveys and total diet studies (market baskets) provide datasets, information and knowledge on arsenic content in foods (D'Amato et al., 2013, Gunderson, 1995, Sirot et al., 2009, Tao and Bolger, 1999, Uneyama et al., 2007, Wong et al., 2013, Ysart et al., 1999). In particularly, the collated data on arsenic in foods by Uneyama et al. (2007) aims to provide a comprehensive comparison data that may be helpful to risk managers and consumers. Additionally, Total Diet Study or Market Baskets conducted by regulatory agencies in countries often include arsenic content in table ready foods.

In September 2012, the US FDA provided data on the assessment of arsenic contents in rice products as part of an ongoing and proactive effort to monitor food safety and address contaminants in food[1]. In July 2013, the US FDA proposed an “action level” of 10 parts per billion (ppb) for inorganic arsenic in apple juice[2]. This level is same as that approved for drinking water by the United States Environmental Protection Agency (USEPA). Consumer Advocates in the United States have also conduct assessments on arsenic in rice products and juices (apple and grape) [3,4].

The determination of the concentration of arsenic in rice varieties (Figure 3) is an active area of research (Banerjee et al., 2013, Gilbert-Diamond et al., 2011, Jackson et al., 2012, Maher et al., 2013, Rintala et al., 2014). Research includes evaluating methods for detection of arsenic in various parts of the rice plant including the grains (Musil et al., 2014, Yamaguchi et al., 2014).

Figure 3. Varieties of rice

Source: (Maher et al., 2013)

---

[1] http://www.fda.gov/NewsEvents/Newsroom/PressAnnouncements/ucm319972.htm
[2] http://www.fda.gov/NewsEvents/Newsroom/PressAnnouncements/ucm360466.htm
[3] http://consumerreports.org/cro/arsenicinfood.htm
[4] http://www.consumerreports.org/cro/magazine/2012/01/arsenic-in-your-juice/index.htm

With the increasing capability to measure arsenic content in foods, there are volumes of varied of continuously generated datasets on arsenic in food groups. Variety datasets are associated with arsenic in foods including arsenic species, country of study or origin, year of study, food groups and method of analysis. Food regulators and researchers need to make sense of data on food contaminants to help consumers make decisions on dietary intake. Consumers should also have the capacity to make sense of consumer reports that are published as static tables.

When compared to tabular presentations, users of visual analytics tools were able to receive more information, see relationships in data more easily, save time and ultimately make more rational decisions (Aragon et al., 2008, Broeksema et al., 2013, Mbah et al., 2013, Savikhin et al., 2008). An information superstructure of arsenic in various foods to permit insightful comparative risk assessment of the diverse and continually expanding data on arsenic in food groups in the context of country of study or origin, year of study, method of analysis and arsenic species.

Several regulatory bodies worldwide including the Joint FAO/WHO Expert Committee on Food Additive, Food Standards Australia New Zealand, World Health Organization and United States Environmental Protection Agency have set various guideline levels for total arsenic, inorganic arsenic and organic arsenic levels in various foods and drinking water (ANZFA, 2001, European Food Safety Authority, 2009, USFDA, 2005, Schmidt, 2012). These regulatory levels are expected to help consumers, risk managers, policy makers and responsible authorities minimize exposure of humans and animals to this toxic element.

Easy to use information technology tools (software and hardware) that harness the human's visual perception abilities to amplify cognition are increasingly recommended to facilitate decision making from data of varying sizes and complexities (Burley and Ashburn, 2010, Thomas and Cook, 2005, Wong and Thomas, 2004). Visual Analytics software and hardware environments facilitate discovery of unknown relationships in datasets through analysis, exploration and mining (Chang et al., 2009, Isokpehi et al., 2012).

Spreadsheet software typically have limitations in visualizing data, creating interactive dashboards, managing unlimited data, and supporting real-time data discovery. [5] There is possibility that future spreadsheet software will include data visualization capabilities. [6]

[5] Tableau vs. Excel http://www.tableausoftware.com/learn/stories/tableau-vs-excel
[6] Excel: Microsoft's best weapon against Tableau and competitors http://www.citeworld.com/article/2114568/consumerization/excel-versus-tableau.html

1.3 Research Goal, Purpose, Hypothesis and Objectives

1.3.1 Goal

The goal of this doctoral research in Environmental Science is to address the need of providing visual analytical decision support tools on arsenic content in various foods. A special emphasis of analysis will be on rice, a staple crop in many countries that has received significant attention regarding arsenic content (Meharg, 2004, Melkonian et al., 2013, Rahman and Hasegawa, 2011, Wei et al., 2013, Williams et al., 2007a, Williams et al., 2007b).

1.3.2 Hypothesis

The hypothesis of this doctoral thesis research is that software enabled visual representation and user interaction facilitated by visual interfaces will help discover hidden relationships between arsenic content and food categories.

1.3.3. Research Objectives

The present study is designed to address the following specific objectives:

1. Provide insightful visual analytic views of compiled data on arsenic in food categories.
2. Categorize table ready foods by arsenic content.
3. Compare arsenic content in rice product categories.
4. Identify informative sentences on arsenic concentrations in rice.

1.4 Rationale for Research Objectives

1.4.1. Objective 1: Provide insightful visual analytic views of compiled data on arsenic in food categories.

The concentrations of arsenic vary in foods making it impractical and impossible to provide regulatory limit for each food. Furthermore, the risks of naturally occurring arsenic in foods have received less attention when compared to drinking water and airborne workplace exposure (Borak and Hosgood, 2007). However, exposure to total and inorganic arsenic from diet is significantly higher than from drinking water (Xue et al., 2010).

Uneyama et al. (2007) have collated data on arsenic content in six food groups (crops, milk/meat/egg, fish, algae, seafood and others) to provide a comprehensive comparison data that may be helpful to risk managers and consumers. The data tables provided by Uneyema et al. (2007) are in a static form which does not allow for efficient human interaction to gain knowledge-building insights on arsenic content from the datasets.

Knowledge-building insights is a form of learning accomplished through visual analytic tools, which builds a relationally semantic knowledge base through a variety of problem-solving and reasoning heuristics (Chang et al., 2009). In summary, knowledge-building insights facilitated by visual analytics tools can enhance decision-making on arsenic in foods by risk managers and consumers.

1.4.2 Objective 2: Categorize table ready foods by arsenic content.

Total Diet Study or market baskets are conducted over a period in several countries with the aim to determine the dietary intake of contaminants and nutrients in foods (Gunderson, 1995). These analytes include selected elements (including radionuclides), pesticides and industrial chemicals (Chen and Gao, 1993, Egan et al., 2002, Gunderson, 1995, Moy, 2013).

Arsenic is one of the elemental analytes determined in United States Food and Drug Administration (FDA) Total Diet Study (TDS). To ensure that estimates of these analytes are realistic of dietary intake, the United States TDS performed analysis on table ready foods (foods prepared as would be consumed) (Egan et al., 2002).

1.4.3 Objective 3: Compare arsenic content in rice product categories.

Arsenic is transported from contaminated ground water into rice grains (Marin et al., 1992, Meharg, 2004, Meharg et al., 2009, Rahman et al., 2009). Inorganic arsenic from cooked rice has been identified as a health risk including in fetal and child development (Gilbert-Diamond et al., 2011, Melkonian et al., 2013).

When arsenic-contaminated rice is consumed, the intestinal tract is the first site of exposure to the rice-derived arsenic that subsequently gets transported to the blood stream and other parts of the body causing a wide range of health problems (Xue et al., 2010).

1.4.4. Objective 4: Identify informative sentences on arsenic concentrations in rice.

With the continuing growth in scientific data, relevant information on the concentration of arsenic in various foods may remain buried in literature. There are over 16,000 scientific abstracts in the PubMed biomedical literature database (www.pubmed.gov) with the MeSH terms arsenic or arsenical. The Comparative Toxicogenomics Database (CTD, http://ctdbase.org) uses PubMed index articles to provide extensive data on chemical:gene interaction including arsenic with relations to genes, diseases (Davis et al., 2013). The focus of CTD has been mainly on toxic relationship with regards to animal systems. This PhD research serves to catalyse the development of visual analytics resources that enable interaction with sentences from PubMed abstracts on concentrations of chemicals in foods.

1.5 Software Used for Visual Analytical Discoveries

The Tableau Software (Tableau Software Inc. Seattle WA) was used as the visual analytics tool to design and share the visual representations and interactions reported in this research (Chabot, 2009, Mackinlay et al., 2007, Murray, 2013). Tableau was selected based on multi-year comparative analysis of it lead in analytical platforms.[7]. In addition to the commercial Desktop Professional version, there are Tableau Reader and Tableau Public versions that are available without subscription costs. It is also possible to publish visualization on the web.

[7] Magic Quadrant for Business Intelligence and Analytics Platforms. http://www.gartner.com/technology/reprints.do?id=1-1QLGACN&ct=140210&st=sb

CHAPTER 2

LITERATURE REVIEW

2.1 Sources of Arsenic

Arsenic is present in the environment as a soil contaminant due to mobilization of the element during natural, anthropological, agricultural and non-agricultural activities (USATSDR, 2007, Matschullat, 2000, Otles and Cagindi, 2010, Smith et al., 1998, Smith et al., 2009). The more toxic of the 2 oxidative states of arsenic is the trivalent state, Arsenite As(III). Inorganic arsenic, is found in water, aquifers, and water generated by environmental occurrences including erosions, weathering, from man-made sources such as smelting and mining (Cadwalader et al., 2011, Razo et al., 2004).

Table 1 below, summarises the different environmental impacting activities that can lead to soil contamination with arsenic (Otles and Cagindi, 2010, Smith et al., 2008). The agricultural activities include use of pesticides, herbicides, fertilizers containing arsenic as well as arsenic contained in livestock feed.

Non-agricultural activities that are sources of contamination of soil by arsenic include the natural weathering of environment which produces regolith, a layer of loose heterogeneous material covering solid rock. Regolith includes dust, volcanic ash and lavas, soil, broken rocks and other related materials) or man-made (smelting, mining, coal and petroleum combustion, by products of industrialization from different factories). Arsenic pollution of the air can also occur from emissions of coal burning for example in China, posing a toxic threat to humans (Millman et al., 2008).

Table 1. Sources of soil contamination by arsenic

Source	Activities
Natural	Regolith originating from weathering and biological activity.
Anthropological	Mining, smelting
Agricultural	Arsenical pesticides, herbicides, fertilizers; arsenic additives in livestock feed
Non-Agricultural	Wood preservatives and coal or petroleum combustion, electronics, industries, pharmaceutical works, galvanizing and ammunition factories

Adapted from (Otles and Cagindi, 2010, Smith et al., 2008)

2.2 Metabolism of Arsenic in the Human Body

The metabolism of inorganic arsenic in the human body is one of the crucial determinants of the toxicity resulting from exposure to inorganic arsenic (Drobna et al., 2010). The metabolic conversion of inorganic arsenic to methylated products is through a multi-step process that results in mono-, di-, and trimethylated arsenicals (Thomas, 2007, Thomas et al., 2007, Thomas et al., 2001). Arsenate (pentavalent arsenic) is reduced to arsenite (trivalent arsenic), which is the preferred substrate for methylation, an oxidative process (Hughes, 2009, Thomas et al., 2007). Glutathione (GSH) donates the electron for the reduction of arsenate to arsinite in aqueous solutions and red blood cells

(Delnomdedieu et al., 1995, Hall et al., 2013, Scott et al., 1993). Arsenate reductases catalyses this reaction in mammalian cells as well as reduction of methylarsonic acid to methylarsonous acid (Radabaugh and Aposhian, 2000, Zakharyan and Aposhian, 1999).

Methylation was thought to be a detoxification process of inorganic arsenic (Dopp et al., 2010, Sumi and Himeno, 2012). However, methylation is now increasingly recognized as a pathway of inorganic arsenic activation. Compared with inorganic arsenicals containing arsenic in the trivalent oxidation state, the methylated forms particularly those in the trivalent oxidation states are more cytotoxic, more genotoxic, and more potent inhibitors of the activities of some enzymes (Muñiz Ortiz et al., 2013, Styblo et al., 2002, Tokar et al., 2013).

The liver is the major site but not the only site for methylation of trivalent inorganic arsenite (iAsIII) in mammals through a folate dependent one-carbon metabolism (D'Amato et al., 2013, Drobna et al., 2010, Gamble et al., 2006, Stamatelos et al., 2011, Vahter, 2002). Transient liver injury was a side effect of arsenic trioxide treatment of promyelocytic leukemia patients (Wang et al., 2013).

In mammals, trivalent arsenic is methylated to trivalent and pentavalent products (Vahter, 2002). The first step of methylation converts inorganic trivalent arsenic As (III) to the trivalent product Methylarsonous acid (MAsIII) and the pentavalent product Methylarsonic acid (MAsV). Subsequent steps in the methylation process yield Dimethylarsonous acid (DMAsIII) and Trimethylarsine (TMAsIII); Dimethylarsonic acid (DMAsV) and Trimethylarsine oxide (TMAs(V)O) for the trivalent and pentavalent forms respectively (Faita et al., 2013, Waters et al., 2004b) (Figure 4). Folate contributes the methyl groups and used in the generation of S-adenosylmethionine (SAM) which has been recognised as the methyl group donor in both in vitro assay and intact animal systems (Drobna et al., 2009, Styblo et al., 1996, Vahter, 2002, Vahter and Marafante, 1987).

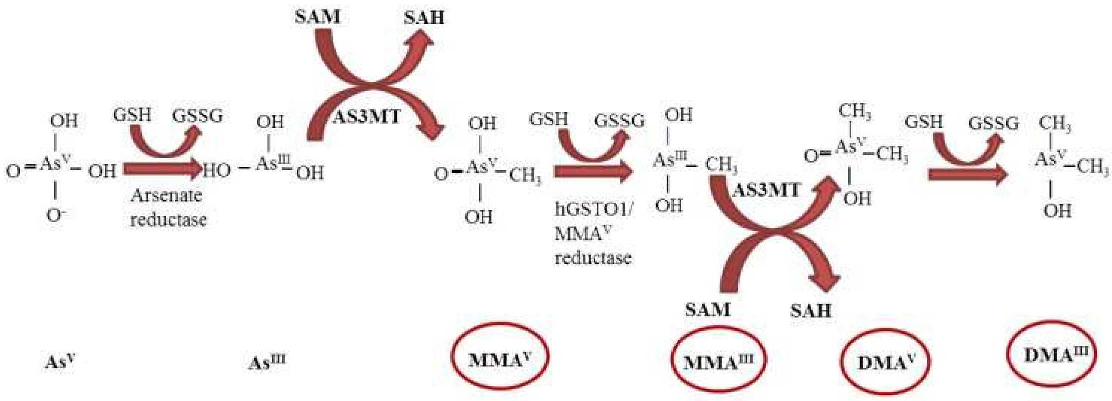


Figure 4. The metabolism pathway of inorganic arsenic showing arsenate reduction to arsenite and methylation to pentavalent and trivalent forms.

Source: (Faita et al., 2013)

In the rat liver, arsenic methyltransferase (AS3MT) or Cyt 19 catalyzes the transfer of the methyl groups from SAM to inorganic arsenite (iAsIII) to generate monomethylarsonic acid MMA(V). After MMA(V) is reduced to monomethylarsonous acid MMA(III), AS3MT can catalyse a second methylation to generate dimethylarsinic acid DMAs(V) (Walton et al., 2003, Waters et al., 2004a, Waters et al., 2004b). In *in vitro* assay systems of rodent liver extracts, conversion of arsenite to methylated metabolites was faster and more extensive than that of arsenate. Likewise, methylarsonous acid was converted to DMAs faster and more extensively than was methylarsonic acid. These processes were found to be inhibited by selenite and mercury (Buchet and Lauwerys, 1985, Styblo et al., 1996).

The major end products of methylation of arsenicals in the human body appears to be DMAs(III) and DMAs(V) (Thomas et al., 2001) and significantly more cytotoxic than inorganic arsenic. These methylated arsenicals enter into the bloodstream and get transported to target tissues and organs including hair, skin, and urine. The methylated trivalent arsenicals, methylarsonous acid MAs(III) and dimethylarsinous DMAsIII differ from their pentavalent counterparts, methylarsonic acid MAsV and dimethylarinic acid DMAsV in that they are more potent cytotoxicants and enzyme inhibitors (Dong et al., 2013, Petrick et al., 2001, Shen et al., 2013).

Three possible reasons exist why the methylated products are thought to be more toxic than the inorganic arsenic. Firstly, methylated products may act as inhibitors, and are more potent inhibitors than are the parent arsenicals, to the action of critical enzymes in the body such as the inhibition of oxidoreductases (e.g thioredoxin reductase, pyruvate dehydrogenase) (Thomas et al., 2001). Secondly, methylated products modulate more efficiently signal transduction pathways that regulate cellular metabolism and survival. (Walton et al., 2003). Thirdly, in a single cell assay system using human blood lymphocytes, both MAs(III) and DMAs(III) were many-fold more potent inducers of DNA damage than was arsenite and arsenate (Mass et al., 2001).

The acute toxicity features to humans of water soluble inorganic arsenic and its metabolites are summarised in Table 2.

Table 2. Selected toxicity features of water soluble inorganic arsenic and its metabolite compounds

Feature	Description
Absorption	Gastrointestinal tract and lungs
Distribution	Liver, kidney, lung, spleen, aorta, and skin
Excretion	Urine at rates as high as 80%
Symptoms of Acute Inorganic Arsenic Poisoning in human	Nausea, anorexia, vomiting, epigastric and abdominal pain, and diarrhea. Dermatitis (exfoliative erythroderma), muscle cramps, cardiac abnormalities, hepatotoxicity, bone marrow suppression and hematologic abnormalities (anemia), vascular lesions, and peripheral neuropathy (motor dysfunction, paresthesia).
Effect of Severe Exposures	Acute encephalopathy, congestive heart failure, stupor, convulsions, paralysis, coma, and death.
General symptoms of chronic arsenic poisoning in human	Weakness, general debility and lassitude, loss of appetite and energy, loss of hair, hoarseness of voice, loss of weight, and mental disorders.
Primary target organs	Skin (hyperpigmentation and hyperkeratosis), nervous system (peripheral neuropathy), and vascular system.
Other symptoms of chronic arsenic poisoning in human	Anemia, diabetes, cancer, leukopenia, hepatomegaly, and portal hypertension.

Adapted from The Risk Assessment Information System: http://rais.ornl.gov/tox/profiles/Arsenic_ragsa.shtml.

2.3 Transport of arsenic in living systems

The entry of arsenic into hepatocytes is controlled by water transport proteins aquaporin (Drobna et al., 2010, McDermott et al., 2010). Aquaglyceroporin channels are transport channels that facilitate bidirectional movement of small neutral solutes such as urea, glycerol. Examples of aquapglyceroporins are the GlpF in *Escherichia coli*, AQP7 and AGP9 from rats and humans (Agre and Kozono, 2003).

The aquaglyeroporins GlpF, AQP7 and AQP9 transport trivalent inoganic arsenic in the form of arsenic trioxide though they differ in selectivity for trivalent arsenicals and transport rates (Liu et al., 2002, Liu et al., 2006). In *Escherichia coli*, GLpF is a channel for conduction of As (III) which produces toxicity (Meng et al., 2004). On the other hand, in another bacteria, rhizobium (***Sinorhizobium meliloti***), another aquaglyceroporin AqpS's actions produces resistance by conducting an efflux of internally generated As(III) (Yang et al., 2005).

The aquaglyceroporin Aquaporin 9 (AQP9) is most expressed in the liver and it transports urea and glycerol in the bile ducts and ductules of the liver. Liu et al. (2006) proposed that AQP9 is essential to arsenic metabolism in the liver because it is the transport channel for As(III) present in the hepatocytes (Figure 5).

As(III) enters the liver down a concentration gradient, then it becomes methylated and reduced. The methylated arsenical products are then transported via AQP9 into the bloodstream out of the cell down a concentration gradient (Liu et al. 2006). Their study also revealed that AQP9 conducts MAs(III) three-fold faster than it does the trivalent inorganic arsenic.

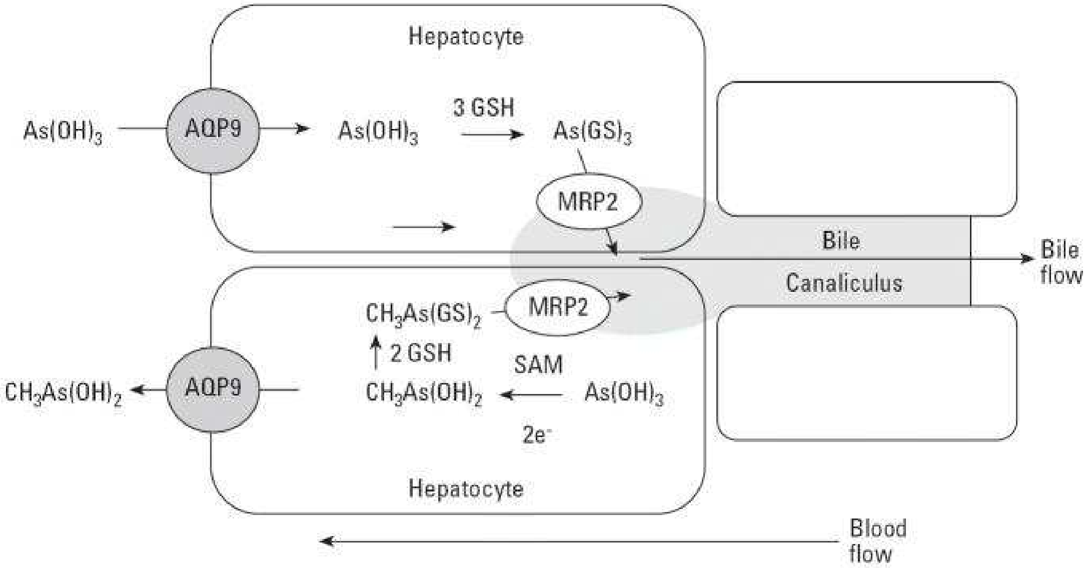


Figure 5. Proposed pathways of trivalent arsenic transport in liver.

Notes: Abbreviations: GSH, glutathione; SAM, S-adenosylmethionine. Trivalent arsenic in the form of As(OH)3 flows down a concentration gradient from blood into hepatocytes through AQP9, which is the major aquaglyceroporin in liver. In the cytosol of the hepatocyte, As(III) can be either glutathionylated or methylated to MAs(V), which is reduced to MAs(III). As(GS)3 is pumped into bile by the multi-drug resistant protein 2 (MRP2), and perhaps by other members of the ABC superfamily of ATPases. Alternatively, As(III) can be methylated and reduced to CH3As(OH)2, which then flows down its concentration gradient via AQP9 into blood. Source: (Liu et al., 2006)

There is also the possibility that nutritional status influences the uptake of trivalent inorganic arsenic and redistribution of methylarsonous acid because the AQP9 expression in rat liver was induced up to 20-fold in the fasting state (Carbrey et al., 2003, Liu et al., 2006). Other variables other than nutritional status, due to the environment or genetic predisposition may bring about inter individual variations in their metabolic capacity of arsenic methylation (Tseng, 2009). These variables include differences in age or sex of individuals (Concha et al., 1998, Del Razo et al., 1997); if pregnant or non-pregnant (Gardner et al., 2012, Gardner et al., 2011); ethnicity (Gomez-Rubio et al., 2012, Vahter et al., 1995); cigarette smoking and alcohol use (Hopenhayn-Rich et al., 1996).

In vitro studies on rat and human hepatocytes revealed a ten-fold variation in the rate of MAs and DMAs formation in different human donors indicating a possible genetic influence on inorganic arsenic methylation (Styblo et al., 1999). Inorganic arsenic and its methylated metabolites are mostly excreted in urine and faeces in four–five days so there is a decreased rate of bioaccumulation (Hughes, 2006). In chronic arsenic ingestion, arsenic accumulates in the liver, kidneys, heart, lungs with smaller amounts in the muscles, nervous system, gastrointestinal tract and spleen (Benramdane et al., 1999).

2.4 Selected Medical, Agricultural and Industrial Applications of Arsenic

In solution at physiologic pH, inorganic trivalent arsenic As(III), is primarily in the form of undissociated acid arsenic trioxide [As(OH)3] (Ramirez-Solis et al., 2004). The anhydrous form of arsenite trioxide (As2O3) is used in the treatment of acute promyelocytic leukemia (APL) (Soignet et al., 1998) by inducing differentiation and apoptotic death in the leukemic cells, thought to be due to an overproduction of reactive oxygen species (Sumi et al., 2010). Data is also available on the molecular responses in human leukemic cell lines as well as enhancement using ascorbic acid (vitamin C) (Yedjou et al., 2010, Yedjou et al., 2009, Yedjou et al., 2006, Yedjou et al., 2008, Yedjou and Tchounwou, 2007, Yedjou and Tchounwou, 2009).

Nitarsone (4-nitrobenzenearsonic acid) is an organoarsenical which is used to prevent histomoniasis (Blackhead) in chicken and turkey that is caused by a flagellated protozoa, *Histomonas meleagridis* (Clarke et al., 2003).

Chromated copper arsenate (CCA), a water-based inorganic arsenic was extensively used to treat lumber for wood preservation purposes (Mercer and Frostick, 2012). The lumber industry voluntary ban the use of the CCA (Hsueh, 2013).

2.5 Possible Mechanisms of Cancer Induction by Inorganic Arsenic

Studies on possible mechanisms of how exposure to arsenite (iAsIII) or arsenate (iAsV) induces cancer suggest that these arsenicals are not direct genotoxicants or mutagens but likely act as a co-carcinogen in combination with a genetic agent (Vogt and Rossman, 2001). In 2010, the International Agency for Research on Cancer (IARC), concluded that exposure to arsenic toxicity in drinking water will cause cancers of the urinary bladder, lung and skin (Straif et al., 2009).

An investigation of American Indians, their low-moderate arsenic exposure and the development of lung, prostate, and pancreas cancer was carried out. (Garcia-Esquinas et al., 2013). Evidence is inadequate to conclude that arsenic toxicity will cause cancers of the kidney, liver or prostate. In Sri Lanka, a possible link of chronic arsenic toxicity with chronic kidney disease of unknown cause (Jayasumana et al., 2013, Jayatilake et al., 2013) (Figure 6).

Figure 6. Abnormal skin pigmentation observed in patients with chronic kidney disease of unknown cause.

Source: (Jayasumana et al., 2013)

Inorganic arsenic modifies signal transduction pathways involved in the regulation of cell growth and proliferation (Simeonova and Luster, 2000). Arsenite modulates the expression and/or DNA binding activities of transcription factors associated with cell proliferation and death including tumour suppressor 53 (p53) (Salazar et al., 2004, Salazar et al., 1997). However, Baastrup et al. (2008), in their study in Denmark, reported that exposure to low levels of arsenic was not associated with cancer rather it may decrease the incidence of non-melanoma skin cancer (Baastrup et al., 2008).

Exposure to inorganic arsenic and the methylated products can cause impaired fetal growth, fetal death in pregnant women and even increased post-birth infant mortality (Vahter, 2008). A summary of emerging understanding of arsenic-induced carcinogenic mechanisms is presented in Figure 7.

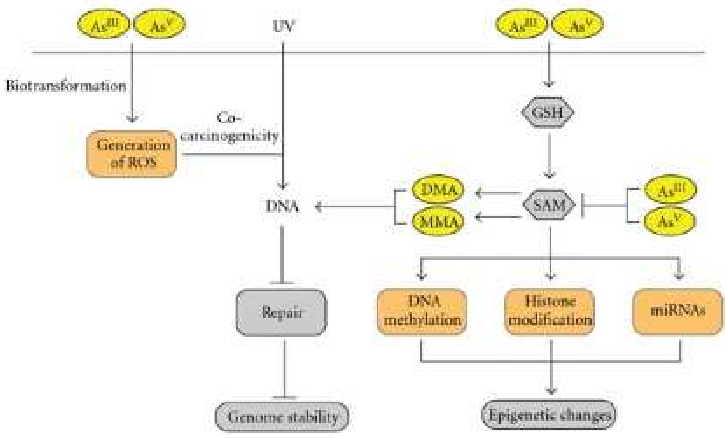


Figure 7. Schematic representation of proposed arsenic-induced carcinogenic mechanisms.

Note: Arsenic can enter cells in both tri- or pentavalent forms (AsIII or AsV). Inside cells, AsV is converted to AsIII, with subsequent methylation to monomethylated (MMA) and dimethylated (DMA) species. The methylation of inorganic arsenic consumes both S-adenosylmethionine (SAM) and glutathione (GSH). Cellular damage derived from arsenic biotransformation can occur through generation of reactive oxygen species (ROS), and through epigenetic mechanisms: changes in DNA methylation patterns (by depletion of cellular pools of methyl group), histone modification, and altered expression of microRNAs (miRNAs). Source: (Martinez et al., 2011)

2.6 Arsenic Toxicity in Young Children

Many research and epidemiologic investigations focus on the effects of arsenic toxicity in adults. However, young children also make up one of the vulnerable groups. This vulnerability of young children to arsenic toxicity can be explained by many physiological reasons and also the behavioural habits they exhibit such as their play habits on the floor, hand or object-mouth habits (Rieuwerts et al., 2006). An investigation of arsenic on the hands of children after playing in playgrounds that use wood treated with chromated copper arsenate (CCA) (Kwon et al., 2004). Comparison of CCA and non-CCA playgrounds revealed a significant difference between the groups ($p < 0.001$). The mean amount of water-soluble arsenic on children's hands from CCA playgrounds was 0.50 μg (range, 0.0078–3.5 μg compared to non-CCA playgrounds, which was 0.095 μg (range, 0.011–0.41 μg).

Extra vigilance is needed to prevent exposure to arsenic from non-dietary objects and from hands. The average mouthing frequency for children aged 6 to <12 months, 12 to <24 months, and 24 to <36 months was determined as 54, 55, and 24 contacts per hour respectively (Beamer et al., 2008). In the United Kingdom (UK), the Department for Environmental, Food and Rural Affairs (DEFRA) and the Environment Agency (EA) have set up a Soil Guideline Value (SGV) for arsenic (Rieuwerts et al., 2006).

Infants and children consume more food per unit body weight than adults making them more vulnerable to the effects of ingested toxic elements. The organs and systems of infants and children are still developing and many of the toxic effects of contaminants may be irreversible and chronic with many effects not becoming apparent till later in life (Dauphine et al., 2011, Smith et al., 2012).

Smith et al. (2012) in a rare in environmental epidemiology that involved a population in Chile exposed to arsenic observed evidence of increased mortality to cancer and non-cancer mortalities among young adults who were in utero or children during the period of high exposure to arsenic (Figure 8).

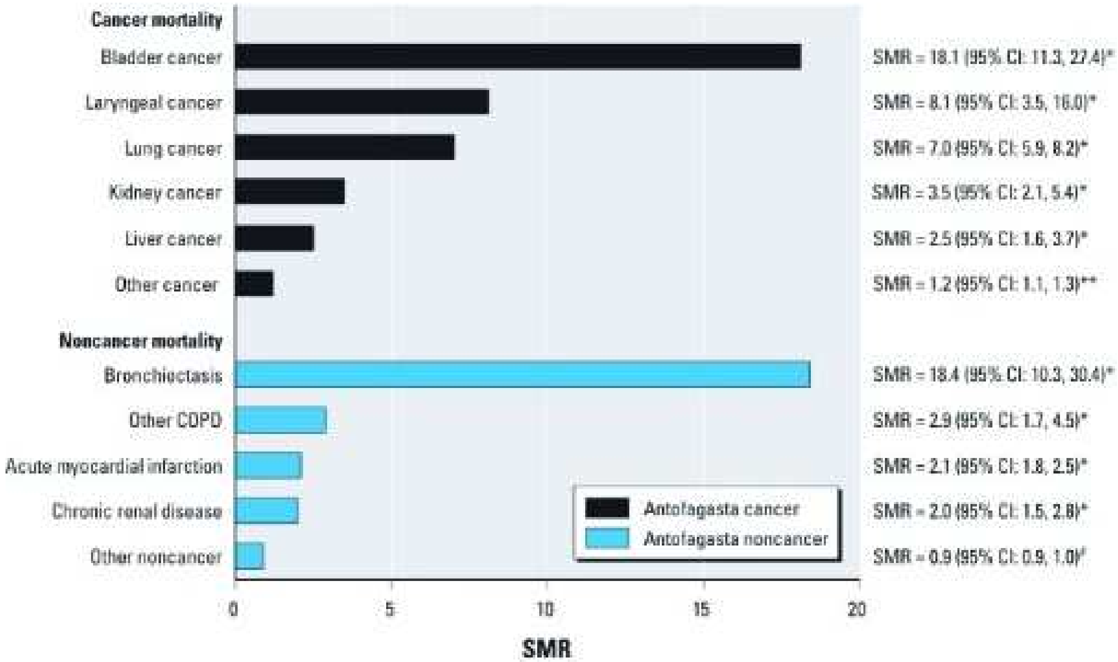


Figure 8. Cancer and noncancer mortality for adults with early-life exposure to arsenic

Notes: Summary of Standardized Mortality Ratios (SMRs) for 30–49-year-old males and females (pooled) who were born in Antofagasta, Chile, combining those born before and during the high-exposure period. *p ≤ 0.001. **p = 0.002. #p = 0.93. Source: (Smith et al., 2012)

In a cross-sectional evaluation of arsenic exposure and cognitive performance of 602 children 6–8 years of age living within 3.5 km of a metallurgic smelter complex in the city of Torreón, Mexico, there was significant inverse association of Urinary Arsenic with several cognition tests (Rosado et al., 2007) (Table 3).

In 1988, the permissible human weekly exposure (PTWI) level for total arsenic set by the Joint FAO/WHO Expert Committee on Food Additives (JECFA) was 0.015mg/kg body weight/week (JEFCA, 1989). However, in the

2011 JECFA, the PTWI for arsenic was withdrawn because it is no longer health protective (JEFCA, 2011).

Table 3. Cognitive scores of children stratified by Urinary Arsenic concentration

Cognitive tests ^	Overall	Children with UAs < 50 µg/L **	Children with UAs > 50 µg/L
Math Achievement Test	31.35 ± 7.50 (3–52)	32.27 ± 7.69 (8–52)	30.57 ± 7.20 (3–49)*
Visual–Spatial Abilities with Figure Design	18.31 ± 5.15 (2–34)	18.88 ± 5.16 (3–34)	17.84 ± 5.08 (2–31)*
WISC-RM Arithmetic Subscale	7.41 ± 3.62 (1–17)	7.26 ± 3.66 (1–17)	7.59 ± 3.57 (1–17)
Peabody Picture Vocabulary Test	103.19 ± 15.65 (55–145)	105.20 ± 16.11 (55–145)	101.67 ± 14.92 (55–140)*
WISC-RM Digit Span Subscale	9.10 ± 3.63 (1–19)	9.46 ± 3.73 (1–19)	8.80 ± 3.55 (2–18)*
Sternberg Memory (correct trials)	12.14 ± 2.94 (4–20)	12.30 ± 3.01 (4–20)	11.98 ± 2.86 (5–20)
Visual Memory Span (correct trials)	2 ± 0.52 (0.69–3.37)	2.03 ± 0.51 (0.69–3.37)	1.97 ± 0.53 (0.69–3.26)
Stimulus Discrimination (correct trials < 19 vs. ≥19)	0.57 ± 0.50 (0–1)	0.63 ± 0.48 (0–1)	0.52 ± 0.50 (0–1)*
WISC-RM Coding Subscale	2.26 ± 0.59 (1–3.71)	2.29 ± 0.58 (1.07–3.71)	2.23 ± 0.61 (1–3.71)
Visual Search (correct minus incorrect minus omitted trials)	5.03 ± 1.51 (1–10.82)	5.23 ± 1.47 (1.14–10.82)	4.84 ± 1.50 (1–8.99)*
Letter Sequencing (correct trials 0 vs. ≥1)	0.48 ± 0.50 (0–1)	0.55 ± 0.50 (0–1)	0.41 ± 0.49 (0–1)*

**mean ± SD (minimum–maximum).
^ WISC-RM; Weschsler Intelligence Scale for Children Revised Mexican Version
*Difference between children with UAs < 50 and children UAs > 50 µg/L is significant at $p<0.05$.
Source: (Rosado et al., 2007)

The European Food Safety Authority (ESFA), states that arsenic has been reported to be a carcinogen in individuals exposed to lower levels than that permitted by JECFA (European Food Safety Authority, 2009) hence ESFA recommends that these levels be re-evaluated, reduced and arsenic levels in different food commodities be determined to help assess dietary exposure. Health authorities have generally found it difficult to establish safety monitoring limits for children's exposure to dietary contaminants because there is an inadequacy of data specific to children's consumption/dietary intakes (Vracko et al., 2009).

2.7 Timeline for Cancer Development from Arsenic Exposure

Arsenic exposure generally progresses from preclinical stage where there are no visible symptoms, detected only in urine or tissue samples, through the clinical or symptomatic and complications stages. The World Health

Organization (WHO) estimated that it takes five to ten years of exposure to arsenic, to full malignancy with cancers of skin and other organs (Otles and Cagindi, 2010, Yoshida et al., 2004).

2.8 Arsenic Concentrations in Various Foods

People in Bangladesh are exposed to arsenic mainly through the food ingestion associated with the consumption of contaminated drinking water and large amounts of rice and other foods (vegetables, daal, fish, milk, chicken and other meats) (Khan et al., 2009). Even in populations where arsenic contamination through water is not a threat, a rice-based diet can contribute a significant amount of arsenic exposure (Meharg et al., 2009, Zhu et al., 2008).

The data on arsenic in six food groups have been collated by Uneyama et al. (2007). In this section, the emphasis is to highlight published articles that have compared arsenic content of selected foods obtained from arsenic-endemic regions of Bangladesh and West Bengal, India with other parts of the world. Furthermore, comparison data on arsenic concentrations in parts of vegetables and grains as well as those found in algae and seafood from Spain and USA respectively are presented.

In Bangladesh, irrigation with underground water has led to increase in the arsenic content of surface soils which then increases the arsenic content of irrigated crops including rice (*Oryza sativa*) (Das et al., 2004). In rice grain samples from 214 households in 25 arsenic-endemic Bangladeshi villages, the total arsenic content ranged from 2 µg/kg to 557 µg/kg dry weight (dw) (Rahman et al., 2009). The arsenic concentrations in control samples obtained from South Australia ranged from 3 µg/kg to 87 µg/kg dw, significantly lower ($p < 0.001$) than those collected in the contaminated areas.

In a survey of arsenic in foodstuffs on sale in the United Kingdom and imported from Bangladesh, the concentration of total arsenic in vegetables from Bangladesh ranged from 5 to 540 µg/kg, with a mean of 54.5 µg/kg (Al Rmalli et al., 2005). Furthermore, the concentration of total arsenic in freshwater fish ranged between 97 and 1318 µg/kg, with a mean value of 350 µg/kg. In the case of freshwater fish, Puti (*Puntius gonionotus*) had a very high arsenic concentration of 1,318 µg/kg with a mean of 580 µg/kg in its dried forms.

The total arsenic concentrations of some selected vegetables including carrots, radish, potatoes, broccoli and cabbage grown in the United Kingdom (UK)/European Union (EU) showed the mean and range of arsenic concentrations to be 24.2 and 5 to 87 µg/kg respectively. The highest concentrations were; 87.2 µg/kg for marrow (a summer squash cultivated in England) and 68.5 µg/kg for cabbage. The comparison of the UK/EU vegetables versus the vegetables imported from Bangladesh, the mean arsenic concentrations are approximately two to three fold higher for Bangladesh.

Roychowdhury et al. (2002) surveyed total arsenic content in food collected in Jalangi and Domkal blocks from the arsenic-affected area of West Bengal, India (Roychowdhury et al., 2002). The food categories surveyed were vegetables (92 and 123 µg/kg), cereals and baked goods (156 and 294 µg/kg) and spices (92 and 201 µg/kg) (mean arsenic concentrations for Jalangi and Domkal blocks respectively).

The arsenic levels in poultry products have been investigated (Ghosh et al., 2012, Kawalek, Lasky, 2013, Lasky et al., 2004, Nachman et al., 2013a, Nachman et al., 2013b). The mean concentration of total arsenic in young chickens was 0.39 ppm, 3- to 4-fold higher than in other poultry and meat (Lasky et al., 2004). In a preliminary study in Bangladesh, the mean arsenic concentrations in egg (wet weight [WW]) was 19.2 ppb (Ghosh et al., 2012). The U.S. FDA in 2011 provided data on various arsenic species present in broilers treated with roxarsone, an inorganic form or arsenic, compared with untreated birds (Kawalek). The geometric mean (GM) of total arsenic in cooked chicken meat samples in a study in the U.S. was 3.0 µg/kg (95% CI: 2.5, 3.6) (Nachman et al., 2013b).

2.9 Arsenic in rice versus other grains

There is a significant emphasis on arsenic levels in rice when compared with wheat which happens to be the second most important food grain worldwide (Zhao et al., 2010). In the wheat grain arsenic is contained mostly in the outer layers of the grain (Zhao et al., 2010). The arsenic concentration in the bran fractions (the outer coverings of the grain) was found to be at least four to five times higher than the content in the white flour fraction (the endosperm).

It is noteworthy that the bran fraction accounts for only 23-29 percent of the total grain weight. This observation corresponds with similar investigations to determine arsenic speciation in rice grain (Lombi et al., 2009, Moore et al., 2010, Sun et al., 2008). Rice bran contains 10 to 20 times higher concentrations of arsenic compared with polished rice and arsenic appears to accumulate in the outer layers of the rice grain.

Wheat and rice differ in the speciation of arsenic in that Zhao et al. (2010) reported that majority of the arsenic extracted from the wheat grain is arsenite, remainder as arsenate but all inorganic arsenic, no methylated arsenic products were extracted. In contrast, if rice is grown anaerobically, in flooded conditions as in paddy fields, most arsenic extracted is inorganic arsenic but there is also a small percentage of arsenic extracted as methylated arsenic, mainly DMA (Meharg et al., 2009, Zavala and Duxbury, 2008, Zavala et al., 2008). This arsenic content is not the case if rice is grown aerobically, it is thought that methylation of the arsenic is a microbial action in the soil (Jia et al., 2013, Xu et al., 2008, Zhao et al., 2013a).

Arsenic concentrations in anatomical parts of vegetables and crops increase in the following order; grain << leaf < stem <<< root. In rice (Liu et al., 2004, Marin et al., 1992), beans (Cobb et al., 2000) elevated concentrations of arsenic in plant roots were observed compared to other plant tissue. Analysis of arsenic concentrations in chard, radish, lettuce and mung beans showed that arsenic accumulated in the following order: root >>> shoot > leaf (Smith et al., 2009).

Furthermore, speciation studies demonstrated that root, shoot and leaf tissue contained only inorganic arsenic with no organic arsenic species identified. Enhanced arsenic shoot assimilation in rice leads to its greater grain levels of arsenic compared with wheat and barley (Williams et al., 2007b). This higher transfer of arsenic in rice can be explained by the increased movement of arsenic in anaerobic versus aerobic soils. The risk of arsenic toxicity to humans via food can therefore be considered in terms of the aerobic versus anaerobic ecosystems.

Estimates of inorganic arsenic in rice have been compiled by the Codex Alimentarius Commission for various countries (CODEX, 2012). The estimates highlight the need to provide estimates for countries in Africa.

Country	Total As		Inorganic As	
	Min-max mg/kg	Mean mg/kg	Min-max mg/kg	Mean mg/kg
Australia	0.05-1.20	0.29	-	-
China	0.08-5.71	0.29	<0.04-0.45	0.13
Japan	0.04-0.43	0.17	0.04-0.37	0.15
EU	0.01-1.98	0.16	0.02-1.88	0.14
UK	0.12-0.47	0.22 (median)	0.06-0.16	0.11 (median)
USA	0.04-0.41	0.21	0.025-0.157	0.091 (different study to min-max values)
Mercosur		0.05-0.13 (parboiled) <0.02-0.03 (polished rice) 0.1 (whole grain)		
Sweden		0.24 (longgrain brown rice) 0.21 (parboiled white rice) 0.1 (white rice)		0.110
Spain		0.197	0.027-0.253	
Slovak Republic		0.158		

Figure 9. Total and inorganic arsenic levels in rice from various countries

Source: (CODEX, 2012)

2.10 Arsenic in seafood

Most arsenic in seafood is organic which is generally considered to be non-toxic (Borak and Hosgood, 2007). In Valencia (Spain), the highest levels of total arsenic in algae food products was obtained from brown algae: Brown algae >>> Red algae >> Green algae (Almela et al., 2002). In a study of heavy metals in commercial fish in New Jersey, USA that some of the fish in the study (Chilean sea bass, croaker, flounder, porgie, and whiting) had arsenic levels of over 1.3 ppm regulatory limit by the USEPA (Burger and Gochfeld, 2013).

2.11 Factors Affecting Arsenic Uptake in Plants

Four geochemical mechanisms of natural arsenic pollution are reductive dissolution, alkali desorption, sulphide oxidation, and geothermal activity (Brammer and Ravenscroft, 2009). Furthermore, many soil factors influence the amount of arsenic available for plant uptake including include redox potential, pH, the contents of organic matter, iron, manganese, aluminium oxides, phosphorus and calcium carbonate, and soil microbes (Mahimairaja et al., 2005). The influence of some of these soil properties and constituents also varies significantly within the year in soils that alternate between anaerobic and aerated conditions. Additional factors are seasonally-flooded soils and irrigated upland soils used for paddy cultivation; the different types of soil; land use and growth seasons (Takahashi et al., 2004, Xu et al., 2008, Yamaguchi et al., 2014).

Plant uptake of arsenic from soils is complicated by a number of factors. In aerated soils used for crops such as wheat, maize and most vegetables, arsenic is present mainly as As(V) and as such is likely to be in the solid phase. Therefore, in such soils, arsenic in groundwater used for irrigation is quickly absorbed by iron hydroxides and becomes largely unavailable to plants. In anaerobic soil conditions such as occur in flooded paddy fields, arsenic is mainly present as As(III) and is absolved in the soil-pore water (the soil solution) (Xu et al., 2008). It is the more readily available to plant roots.

Emerging evidence indicate that dimethylarsinic acid (DMA), a methylated species of arsenic, is derived from the soil and produced by soil microorganisms through the action of the protein for arsenic methylation (arsM) (Figure 10) (Zhao et al., 2013a, Zhao et al., 2013b). Rice, tomato and clover plants lack the ability to methylate inorganic arsenic (Lomax et al., 2012). Furthermore, microbial methylation of arsenic is increased in the soil during flooding and addition of organic matter (Jia et al., 2013, Zhao et al., 2013a, Zhao et al., 2013b). Compared with inorganic arsenic, the uptake of DMA by rice plant roots is less efficient but their transport to the grain is much more efficient (Zhao et al., 2013b). Silicon decreases the accumulation of methylated arsenic species in rice (Liu et al., 2013).

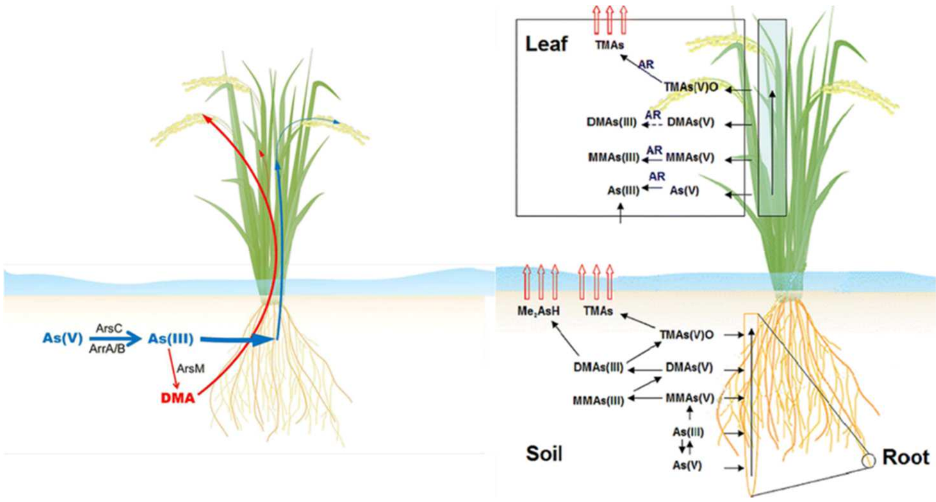


Figure 10. Arsenic Uptake and Volatilization from Rice Plants and Paddy Soil

Notes:
Uptake: Compared with iAs, methylated As species are taken up by rice roots less efficiently but are transported to the grain much more efficiently, which may be an important factor responsible for the spikelet sterility disorder (straight-head disease) in rice. Source: (Zhao et al., 2013a)
Volatilization: In the axenic system, uptake of As species into rice roots was in the order of arsenate (As(V)) > monomethylarsonic acid (MMAs(V)) > dimethylarsinic acid (DMAs(V)) > trimethylarsine oxide (TMAs(V)O), but the order of the root-to-shoot transport index (Ti) was reverse. Also, volatilization of trimethylarsine (TMAs) from rice plants was detected when plants were treated with TMAs(V)O but not with As(V), DMAs(V), or MMAs(V). Source: (Jia et al., 2012)

2.12 Arsenic Estimates in Total Diet Study

The Total Diet Study (TDS) is a continuous market basket-type survey in which foods representing the average diet of a population or country are purchased, prepared and combined into groups of similar foods for analysis of non-nutrients (such as contaminants and chemical residues) and nutrients (such as sodium and manganese) (Moy, 2013, Rose et al., 2010). Country-based TDS is a risk assessment conducted periodically (e.g. annually and every five years) with a primary purpose of measuring the average amount of each chemical ingested by different age/sex groups living in a country (Betsy et al., 2012, Moy, 2013, Rose et al., 2010). Figure 11 shows a 2005 map produced by the World Health Organization indicating countries where at least one Total Diet Study has taken place or planned.

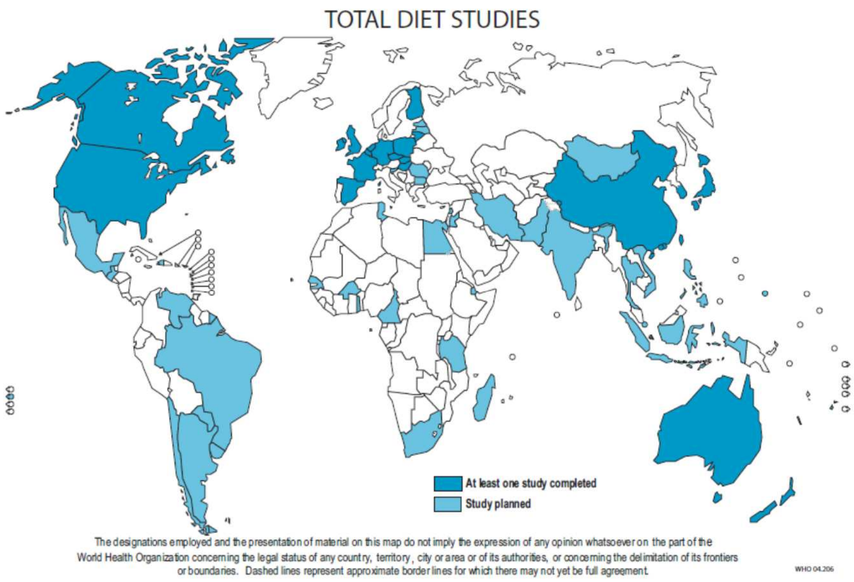


Figure 11. Global distribution of Total Diet Studies

Source: http://www.who.int/foodsafety/chem/TDS_recipe_2005_en.pdf

The TDS and other food consumption surveys provide estimates of exposures to elements including arsenic (Egan et al., 2007). These estimates can be compared to Health-Based Guidance Values (HBGV) provided by governmental food safety organizations including the European Food Safety Authority and the Joint FAO/WHO Expert Committee on Food Additives (JECFA) and United States Food and Drug Administration (D'Amato et al., 2013).

In the United States, the TDS was initiated based on concerns in the 1950s about dietary exposure to radionuclide fallout from nuclear weapons testing and the residues of chemical pesticides (Egan, 2013). Total diet studies are recommended by the World Health Organization as a cost-effective measure to ensure the safety of foods consumed by the public.

A Total Diet Study produces results on (i) the dietary exposures; and (ii) the analytical results of the foods that are sampled (Flynn, 2013). Both types of results need to be communicated to stakeholders including government, industry, academia, consumer organizations and consumers. One representation of the key results is multidimensional tables with chemicals in the columns and foods in rows (Figure 12). For example in the United States TDS, 280 foods are analyzed for 16 elements, selected chemical residues and pesticides (Flynn, 2013). Results may also be presented as tables showing analytes covered by the TDS for each food (Figure 13) as in the Food Safety Authority of Ireland (FSAI) Report on TDS for the period 2001 to 2005 (FSAI, 2011).

Table 3a. Concentrations (in milligrams per kilogram) of aluminium (Al), antimony (Sb), arsenic (As), barium (Ba), bismuth (Bi), cadmium (Cd), chromium (Cr), copper (Cu), germanium (Ge), indium (In) and lead (Pb) in the 20 food groups of the 2006 UK Total Diet Study

Food Group	Al	Sb	As		Ba	Bi	Cd	Cr	Cu	Ge	In	Pb
			Inorganic	Total								
Bread	3.59	(0.0014)	<0.01	<0.005	0.81	<0.001	0.023	<0.02	1.66	<0.002	<0.02	(0.011)
Misc. Cereal	17.5	0.0020	(0.012)	0.018	0.74	<0.001	0.021	(0.03)	2.21	<0.002	<0.02	(0.007)
Carcase Meat	0.24	(0.0008)	<0.01	(0.006)	(0.03)	<0.0005	<0.003	(0.03)	1.44	<0.001	<0.01	<0.003
Offal	0.22	(0.0008)	<0.01	(0.008)	0.09	<0.0005	0.084	<0.01	52.5	(0.002)	<0.01	0.065
Meat Product	2.50	0.0099	<0.01	(0.005)	0.33	<0.0005	(0.007)	0.037	1.16	(0.001)	<0.01	(0.005)
Poultry	0.20	(0.0008)	<0.01	0.022	(0.03)	<0.0005	<0.003	<0.01	0.72	<0.001	<0.01	<0.003
Fish	0.81	0.0026	(0.015)	3.99	0.14	(0.0006)	0.015	0.04	0.91	<0.0007	<0.007	(0.004)
Oils & Fats	0.27	<0.0005	<0.01	<0.005	<0.04	(0.001)	<0.005	0.02	(0.08)	<0.002	<0.02	<0.006
Eggs	<0.03	<0.0003	<0.01	<0.003	0.33	<0.0005	<0.003	0.01	0.57	<0.001	<0.01	<0.003
Sugar & Preserves	2.73	0.0044	<0.01	(0.009)	0.49	0.005	(0.006)	0.08	1.80	<0.002	<0.02	<0.006
Green Vegetables	1.12	0.0005	<0.01	0.004	0.465	(0.0005)	0.006	(0.008)	0.580	<0.0003	<0.003	0.004
Potatoes	0.98	(0.0004)	<0.01	(0.005)	0.17	(0.0005)	0.028	0.031	1.12	<0.0007	<0.007	(0.003)
Other Vegetables	2.84	0.0055	<0.01	0.005	0.533	(0.0004)	0.007	0.024	0.808	<0.0004	<0.004	0.013
Canned Vegetables	1.02	0.0005	<0.01	(0.001)	0.249	0.0009	0.006	0.039	1.29	<0.0004	0.096	0.006
Fresh Fruits	0.48	0.0004	<0.01	(0.001)	0.422	(0.0003)	<0.001	(0.007)	0.786	<0.0003	<0.003	(0.002)
Fruit Products	1.17	0.0004	<0.01	(0.003)	0.212	(0.0003)	0.004	0.017	0.544	<0.0003	0.031	0.007
Beverages	1.49	0.0004	<0.01	<0.001	0.036	<0.0002	<0.001	<0.003	0.074	<0.0003	<0.003	(0.001)
Milk	(0.01)	<0.0001	<0.01	<0.001	0.070	0.0020	<0.001	<0.003	0.050	<0.0003	<0.003	(0.001)
Dairy Products	0.50	(0.0004)	<0.01	<0.003	0.22	0.0064	<0.003	(0.01)	0.33	<0.001	<0.01	<0.003
Nuts	3.81	(0.0007)	<0.01	(0.007)	131	<0.001	0.065	(0.03)	9.15	<0.002	<0.02	<0.006

Brackets indicate the measured values are below the LOQ; LODs and LOQs for a given element will vary according to the weight of sample taken.

Figure 12. Sample tabular representation from a Total Diet Study

Currently the information from most Total Diet Studies are available in electronic formats (such as portable document format) that do not allow for dynamic interaction with and knowledge building from the data. These static presentations of the TDS results can be significantly improved using visual analytics. Visual representation and interaction with the data from TDS can provide a means to communicating the estimates of exposures and results to stakeholders.

The next section of this doctoral thesis chapter provides an overview of visual analytics as a tool for (i) analytical reasoning; (ii) visual representations and interaction; (iii) data representations and transformations; and (iv) the production, presentation and dissemination of results of analysis.

ANNEX 4: ANALYTES COVERED BY TDS

Analyte Abbreviations			
Al = Aluminium	Pb = Lead	NO2 = Nitrite	FBs = Fumonisins (FB1, FB2)
As = Arsenic	Hg = Mercury	Benz. = Benzoate	AM1 = AflatoxinM1
As(i) = Inorg. Arsenic	Se = Selenium	Sorb. = Sorbate	Pat = Patulin
Cd = Cadmium	Na = Sodium	SO3- = Sulfite	AFs = Aflatoxins (B1, B2, G1, G2, Total)
Cr = Chromium	I = Iodine	AA = Acrylamide	Trich = Trichothecenes
Sn = Tin	F- = Fluoride	PAHs = PAHs (15 SCF)	Pest = Pesticides
Sr = Strontium	NO3 = Nitrate	OTA = Ochratoxin A	

Food	Analytes
White flour	As, Cd, Cr, Sr, Pb, Hg, Se, Na, I, OTA, FBs, Trich, AFs,
Wholemeal flour	As, Cd, Cr, Sr, Pb, Hg, Se, Na, I, OTA, FBs, Trich, AFs,
White bread and rolls	As, As(i), Cd, Cr, , Sr, Pb, Hg, Se, Na, I, F-, NO3-, NO2-, Benz., Sorb., SO3--, AA, PAHs, OTA, FBs, Trich, AFs,
Brown bread and brown rolls	As, As(i), Cd, Cr, Sr, Pb, Hg, Se, Na, I, F-, NO3-, NO2-, Benz., Sorb., SO3--, AA, PAHs, OTA, FBs, Trich, AFs,
Plain biscuits	As, As(i), Cd, Cr, Sr, Pb, Hg, Se, Na, I, F-, NO3-, NO2-, Benz., Sorb., SO3--, AA, PAHs, OTA, FBs, Trich, AFs,
Chocolate biscuits	As, As(i), Cd, Cr, Sr, Pb, Hg, Se, Na, I, F-, NO3-, NO2-, Benz., Sorb., SO3--, AA, PAHs, OTA, FBs, Trich, AFs,
Cakes	As, As(i), Cd, Cr, Sr, Pb, Hg, Se, Na, I, F-, NO3-, NO2-, Benz., Sorb., SO3--, AA, PAHs, OTA, FBs, Trich, AFs,

Figure 13.Sample tabular representation from Total Diet Study conducted in Ireland

Source: Total Diet Report at http://www.fsai.ie/

2.13 Goal of Visual Analytics

Analysis is both an art and a science. The goal of analysis is to make judgments about an issue or larger questions. The focus areas of visual analytics are summarized in Table 4. The perception is that visual analytic techniques are developed for massive datasets and complex problems. Chabot (2009) argues that visual analytics techniques are for everyday use for both large and small multidimensional data as well as for answering simple and complex questions (Chabot, 2009).

In addition, it is not always about finding hidden insights about the data, but exploring, cleaning, gaining confidence in, summarizing, pursing inconclusive paths, confirming facts and presenting findings about the data. In other words, visual analytics is an iterative process that involves collecting information, data pre-processing, knowledge representation, interaction, and decision-making (Keim et al. 2006). In summary, the goal of visual analytics tools is to enable people apply computing operations to data by interacting directly with visual representations.

Table 4: Visual Analytics Focus Area Techniques

Focus Area	Function to Users
Analytical reasoning	Obtain deep insights into the data at hand that will directly support assessment, planning and decision making
Visual representations and interaction	See, explore, and understand large amounts of information at once
Data representations and transformations	Convert data which may previously have appeared in all types of conflicting and dynamic into ways that support its visualization and analysis
Support the production, presentation and dissemination of results of analysis	Communicate the information in the appropriate context to a variety of audience

Adapted from Thomas and Cook (2005).

2.14 Visual Analytics as an Integrated Approach

Analytically important data are buried in vast streams of all types. Raw data, are rarely appropriate for direct analysis hence visual analytics must bring all relevant data into a single consistent analytical context, regardless of the form in which the information began, to support analysis and discovery (Thomas and Cook, 2005).

Information visualization draws on the intellectual history of several traditions, including computer graphics, human-computer interaction, cognitive psychology, semantics, graphic design, statistical graphics, cartography, and art. According to Andrienko *et al.* (2008) “In a strict sense, visualisation is representation of data in a visual form, i.e. creating various pictures from data: graphs, plots, diagrams, maps, etc. For this purpose, items of data are translated into graphical features, such as positions within a display, colours, sizes, or shapes.”

Furthermore, Andrienko *et al.* (2008) states, “the main goal of data exploration is detecting patterns and relationships in the data”. For these visualisation tools to be effective for the exploration of data, they should allow visualisation of all the data in a single interface, allow users to interact with and manipulate the data to access various details hidden in the data (Andrienko et al., 2007).

Visual analytics is thus more than visualization but is an interdisciplinary field of research with a scope involving many fields including knowledge discovery, information analytics amongst others. It draws strength from these other fields in order to gain insight into data of various sizes and complexity. It is also an integrated approach combining fields such as visualization, human factors and data analysis, which in turn integrate different methods as shown in Figure 14.

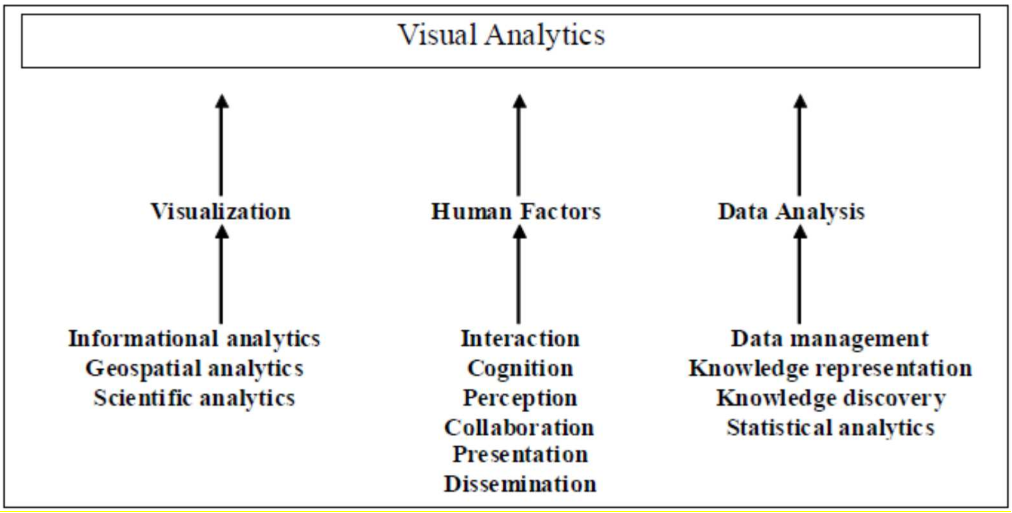


Figure 14. Visual Analytics as an integrated approach.

2.15 "Insight" in Visual Analytics

"Insight" in visual analytics has quite a few definitions but none is commonly accepted as a definition by the community of visualization (Plaisant et al., 2008, Saraiya et al., 2006). Researchers in the area of cognitive neuroscience define insight as that 'eureka' moment, when a person moves from the point of not knowing the solution to a problem to the point of knowing. It is detectable by measuring the neural activity using an Electroencephalography (EEG) or functional Magnetic Resonance Imaging (fMRI) (Lehrer, 2008). This is a spontaneous moment (Mai et al., 2004) and often the thought process leading up to this solution occurred in a subconscious state (Bowden et al., 2005).

The community of visualization defines insight as "the gaining of knowledge about a data after interactively visualizing and exploring it". It is thus knowledge-building and not spontaneous. They also define insight as "new information discovered that could bring to light previously unknown relationships in the data" (Chang et al., 2009).

To measure the amount of knowledge-building insight, the methods used to gather the knowledge are evaluated as well as studies to measure the amount of knowledge gained by a user. Thus, in visual analytics and information visualization, insight can be discovered, gained or provided whereas in cognitive science, insight is experienced making it an event and not a substance. It has been proposed that spontaneous insight in fact comes from knowledge about a problem and each spontaneous insight can open up new directions for more knowledge-building.

2.16 Visual Analytics Tools

Visual analytics tool has been defined as a software or application, through which users may interact with and explore data using the visual interface, and arrive at insights that are not obvious from staring the underlying raw data [8]. The number of visual analytics tools is growing. A collection of educational resources on visual analytics can be found at the Visual Analytics Community website http://www.vacommunity.org/Education+Resources.

A Visual Analytics Toolkit has been defined as both automated analysis functionality and information visualization techniques in an integrated programming library for stand-alone applications or plugin development (Harger and Crossno, 2012). Table 5 and Table 6 respectively presents list and descriptions of selected toolkits and tools for visual analytics. Figure 15 and Figure 16 are comparisons of visual analytics tools and open source visual analytics toolkits respectively.

[8] http://www.inetsoft.com/literature/Visual_Analytics_for_the_Masses.pdf

Tableau software, which is used in this research is briefly described here. Tableau Software is a commercial application package that allows for data to be explored, visualized, analyzed and the results shared (Mackinlay et al., 2007). In addition to the commercial Desktop Professional version, there are Tableau Reader and Tableau Public versions that are available without subscription costs. Gartner Incorporated reported that Tableau continues as a leader in Magic Quadrant for Business Intelligence and Analytics Platforms.[9]

The Magic Quadrant consist of 27 analytics platforms which where compared using 17 categories group into Information Delivery, Analysis and Integration. The Gartner report identified the following capabilities of Tableau Software: highly intuitive, visual-based data discovery, dashboarding, and data mashup without the need for extensive skills or training in business intelligence platform. Since potential users of the visual representations from this research would have minimal or no training in visual analytics tools, the visual representations of data in this research was done using Tableau Software. A visualization of Table 7 using Tableau Software is presented in Figure 17. A table of 30+ data visualization and analysis tools is available at the website of Computer World Magazine [10] (Machilis, 2013).

In the area of environmental health risk assessment, the knowledge explosion available in scientific literature has made it impossible to manually extract facts from abstracts or entire articles (Chen et al., 2014, Klassen et al., 2010, Choi et al., 2013). An added analytical task capability will be to discover

---

[9] http://www.gartner.com/technology/reprints.do?id=1-1QLGACN&ct=140210&st=sb

[10] http://www.computerworld.com/s/article/9214755/Chart_and_image_gallery_30_free_tools_for_data_visualization_and_analysis

and integrate quantitative and qualitative data from fragments of text such as sentences. PubMed literature database (www.pubmed.gov) provides abstracts that can be segmented into sentences for subsequent analytical tasks including natural language queries. The fourth objective of this project addresses new analytical tasks involving data discovery from scientific text.

Table 5: Selected Toolkits for Visual Analytics

Toolkit	Description and Website
Prefuse	A set of software tools for creating rich interactive data visualizations. The original prefuse toolkit provides a visualization framework for the Java programming language. The prefuse flare toolkit provides visualization and animation tools for ActionScript and the Adobe Flash Player. http://prefuse.org/
Flare	An ActionScript library for creating visualizations that run in the Adobe Flash Player. From basic charts and graphs to complex interactive graphics, the toolkit supports data management, visual encoding, animation, and interaction techniques. http://flare.prefuse.org/
Gephi	An open-source, Java toolkit for visualizing and analysis network/graph data. Have many layout algorithms and social network analysis algorithms (centralities) implemented. Support networks/graphs up to half million edges. http://gephi.org/
GeoViz Toolkit	The GeoViz Toolkit supports systematic analysis of spatial, temporal, and attribute data sets. http://www.geovista.psu.edu/geoviztoolkit/
Improvise	A fully-implemented Java software architecture and user interface that enables users to build and browse highly-coordinated visualizations interactively. http://www.cs.ou.edu/~weaver/improvise/index.html
GUESS	Graph Exploration System - an exploratory data analysis and visualization tool for graphs and networks. http://graphexploration.cond.org/
JUNG	Java Universal Network/Graph Framework - A software library that provides a common and extendible language for the modeling, analysis, and visualization of data that can be represented as a graph or network. http://jung.sourceforge.net/index.html
NWB Network Workbench	A Large-Scale Network Analysis, Modeling and Visualization Toolkit for Biomedical, Social Science and Physics Research. http://nwb.cns.iu.edu/

Table 6. Selected Tools for Visual Analytics

Tool	Description and Website
NodeXL	A template for Excel 2007 and 2010 that lets you enter a network edge list, click a button, and see the network graph in the Excel window. You can easily customize the graph's appearance; zoom, scale and pan the graph; dynamically filter vertices and edges; alter the graph's layout; find clusters of related vertices; and calculate a set of graph metrics. Networks can be imported from and exported to a variety of data formats, and built-in connections for getting networks from Twitter, Flickr, YouTube, and your local email are provided. http://nodexl.codeplex.com/
Jigsaw	A visual analytics system to help analysts and researchers explore, analyze, and make sense of document collections in order to reach timely, accurate understandings of the larger stories and important concepts embedded in textual reports. It provides visualizations of different aspects of the documents, presenting the identifiable important entities (people, places, organizations, etc.) and their direct or indirect connections helping analysts explore relationships and connections among the entities. http://www.cc.gatech.edu/gvu/ii/jigsaw/
Tableau	Browser-based analytics and data visualization. Tableau Public is for anyone who wants to tell stories with interactive data on the web. It's delivered as a service which allows you to be up and running overnight. With Tableau Public you can create amazing interactive visuals and publish them quickly, without the help of programmers or IT. http://www.tableausoftware.com/
IN-SPIRE	A discovery tool that integrates information visualization with query and other interactive capabilities. It automatically conveys the gist of large sets of unformatted text documents such as technical reports, web data, newswire feeds and message traffic. It can handle real-time data by adding new documents as they arrive and processes foreign language data providing robust support for translation. By clustering similar documents together, this Windows-based software unveils common themes and reveals hidden relationships. http://in-spire.pnnl.gov/
R	R is a language and environment for statistical computing and graphics. It is a GNU project similar to S developed at Bell Labs. R provides a wide variety of statistical (linear and nonlinear modelling, classical statistical tests, time-series analysis, classification, clustering, ...) and graphical techniques, and is highly extensible. One of R's strengths is the ease with which well-designed publication-quality plots can be produced, including mathematical symbols and formulae where needed. http://www.r-project.org/
KNIME	A modular data exploration platform that enables the user to visually create data flows (often referred to as pipelines), selectively execute some or all analysis steps, and later investigate the results through interactive views on data and models. http://www.knime.org/
GGobi	An open source visualization program for exploring high-dimensional data. It provides highly dynamic and interactive graphics such as tours, as well as familiar graphics such as the scatterplot, barchart and parallel coordinates plots. Plots are interactive and linked with brushing and identification. http://www.ggobi.org/
Many Eyes	A project designed to democratize visualization by harnessing the power of human visual intelligence to find patterns and enabling anyone on the internet to publish powerful interactive visualizations and start their own data conversations. http://www-958.ibm.com/software/data/cognos/manyeyes/
Weave	(Web-based Analysis and Visualization Environment), data visualization platform, free to public and nonprofit users in late 2010, from U Mass Lowell, see also OpenIndicators http://oicweave.org/

Source: Gartner (February 2014)

Figure 15 Comparison of visual analytics tools by ability to execute and completeness of vision.

Additional information on the comparison is available at
http://www.gartner.com/technology/reprints.do?id=1-1QLGACN&ct=140210&st=sb

	Bar Chart	Line Chart	Scatterplot	Box Plot	Pie Chart	Contour Plot	Stacked Bar Chart	Stacked Area Chart	Parallel Coordi-nates
Axiis	✓	✓	✓		✓		✓	✓	
birdeye	✓	✓	✓		✓		✓	✓	
Flare	✓	✓	✓					✓	
Gephi									
Google Vis	✓	✓	✓	✓	✓		✓	✓	✓
GraphViz									
Improvise[a]	✓	✓	✓					✓	
IVTK	✓	✓	✓						✓
JIT							✓	✓	
JFreeChart	✓	✓	✓	✓	✓		✓	✓	
JGraph									
JUNG									
NetworkX									
Prefuse	✓	✓	✓					✓	
Protovis	✓	✓	✓	✓	✓		✓	✓	✓
R	✓	✓	✓	✓	✓	✓	✓		
Titan	✓	✓	✓		✓				✓
Tulip			✓						✓
VisAD		✓	✓			✓			
WilmaScope									
Zest									

[a] Uses Prefuse for visualization

Figure 16 Comparison of visual analytics toolkits by visualization functionality for tabular data

Source: (Harger and Crossno, 2012)

Table 7. Total arsenic content (mg/kg) of selected foods from a study in New Zealand

Food	Brand 1	Brand 2	Brand 3	Brand 4
Apple-based juice	0.001	< 0.001	0.002	0.003
Apricot, canned	< 0.002	< 0.002	< 0.002	< 0.002
Beer	0.003	< 0.001	0.001	0.001
Biscuit, chocolate	< 0.010	< 0.010	< 0.010	< 0.010
Biscuit, cracker	0.010	0.020	0.020	< 0.010
Bran flake cereal, mixed	0.020	< 0.010	0.020	< 0.010
Caffeinated beverage	< 0.001	< 0.001	< 0.001	< 0.001
Chicken	0.009	0.011	0.010	0.010
Chocolate beverage	0.001	< 0.001	< 0.001	< 0.001
Fish fingers	0.873	0.727	0.485	0.790
Fish, canned	0.610	0.572	1.090	0.866
Infant weaning food, cereal based	0.003	0.002	0.011	0.012
Infant weaning food, custard/fruit dish	0.043	0.005	0.009	0.011
Infant weaning food, savoury	0.025	< 0.002	0.003	0.007
Muesli	0.010	< 0.010	0.010	< 0.010
Noodles, instant	0.003	0.005	< 0.002	< 0.002
Oats, rolled	< 0.002	0.004	< 0.002	< 0.002
Oil	< 0.010	< 0.010	0.020	< 0.010
Pasta, dried	0.003	< 0.002	< 0.002	0.003
Peaches, canned	0.002	< 0.002	< 0.002	< 0.002
Prunes	< 0.002	< 0.002	< 0.002	0.003
Raisin/sultana	0.007	0.017	0.008	0.021
Rice, white	0.101	0.039	0.031	0.050
Snack bars	< 0.010	0.010	0.020	< 0.010
Soy milk	0.004	0.003	0.002	0.094
Spaghetti in sauce, canned	< 0.002	< 0.002	0.032	< 0.002
Wheatbix	< 0.010	< 0.010	0.020	< 0.010
Wine, still red	0.010	0.006	0.004	0.004
Wine, still white	0.004	0.004	0.007	0.009
Yeast extract	0.237	0.148		

Source: http://www.nzfsa.govt.nz/science/research-projects/total-diet-survey/reports/quarter-2/quarter-2-nztds.pdf

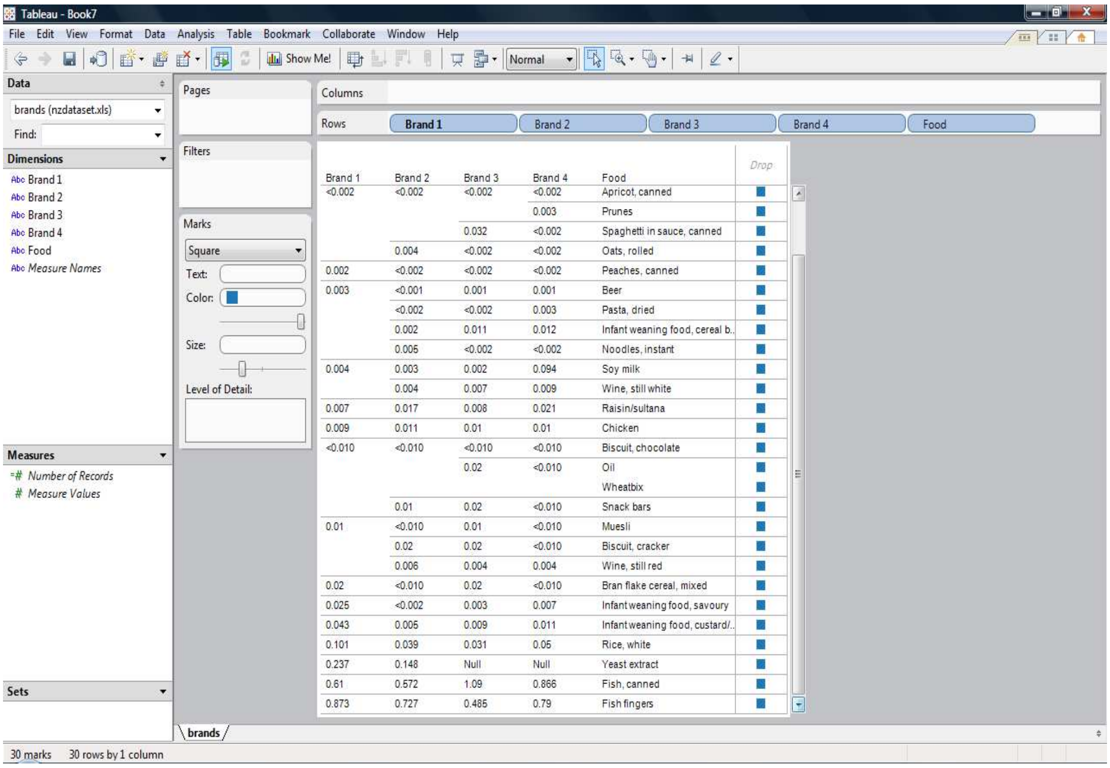


Figure 17. Screenshot of visual analytics interface for grouping arsenic content (mg/kg) of foods from a study in New Zealand.

Notes: Visual Analytics process revealed relationship between Oil and Wheatbix.

2.17 Measuring Insight from Visualization Tools

Saraiya et al. (2006) citing Spence (2001) and Card et al. (1999) reported that the main goal of a visual analytics tool is to enable users gain insight into the data being visualized (Card et al., 1999, Saraiya et al., 2006). Insight is a unit of discovery made during an individual observation of data by a participant (Saraiya et al., 2005). Generally studies done to evaluate effectiveness of visualization tools do so by measuring and analyzing the performance time and accuracy of study participants' responses to pre-determined tasks. Variables measured in these studies include accuracy by measuring precision, error rates, number of correct and incorrect responses and performance assessed by measuring time taken to complete predefined tasks (Chen and Czerwinski, 2000, Kobsa, 2001, Saraiya et al., 2005, Saraiya et al., 2006). To overcome the limitation of these methods of evaluation of insight, Saraiya *et al.* (2006) proposed that a longitudinal study provides a true evaluation of insight since long-time insight is gained by users spending more time studying the data especially if they had no prior knowledge of the visualization tool used or the experimental data.

For users to effectively gain insights into data using visual analytics tools in an efficient and enjoyable process, these tools must have a combination of visual representations and interaction mechanisms (Saraiya et al., 2006). Participants in their study found the most exciting insights after about one and half months of data analysis and spending enough time to learn about the software. Finally, the method of visualization must be appropriate for the data type.

2.18 Literature Curation and Extraction of Facts from Biomedical Literature

Literature curation is identifying scientific data in literature and depositing in a database appropriately, but it requires expertise and it takes time (Alex et al., 2008). Facts can be extracted from biomedical literature in two basic ways and made accessible. One method is by manual curation of these facts. This method is highly accurate but time consuming and cannot keep up with the increasing biomedical data. Also, pertinent information may be left out due to human oversight (Muller et al., 2004). Manual annotation is a subjective process, being dependent on each annotator's scientific background. Sometimes inter-annotator agreement can be as low as 39% (Camon et al., 2005).

Natural Language Processing (NLP) is the scientific field dedicated to training computers with the right knowledge to understand text. Text Mining is an aspect of NLP (Rodriguez-Esteban, 2009). In contrast, automated methods derived from the Natural Language Processing and Information retrieval do not exhibit the problems found with manual annotation. They are not time consuming and they can be applied to large volumes of text.

Automation allows for the systems to be updated once there are new concepts. Limitation of automation is that the quality is not as high as manual curation. The knowledge extraction aspect is still prone to error and needs to be confirmed by a curator (Winnenburg et al., 2008). Text mining, the automated method of information retrieval, scales well with increasing data but can be error prone due to complexities of natural language.

Wiegers *et al.* (2009) conducted a study to develop a text-mining tool from existing components that can improve curation efficiency as well as increase data coverage for the Comparative Toxicogenomics Database in a single workflow (Wiegers et al., 2009). Text mining was found to be a useful aid to manual curation by increasing productivity and increasing quality of data curated. In particular, Text mining tools were also able to prioritize effectively the relevant articles for manual curation.

The focus of this thesis is on the curation of sentences in PubMed abstracts for concentration of arsenic in foods. The use case is arsenic concentrations in rice reported in in PubMed abstracts available at http://www.pubmed.gov/. According to the Oxford Dictionary, a sentence is defined as “a set of words that is complete in itself, typically containing a subject and predicate, conveying a statement, question, exclamation, or command, and consisting of a main clause and sometimes one or more subordinate clauses”.

In the context of scientific literature, Wilbur et al (2006) believes that “an important first step towards a more accurate information extraction and retrieval, lies in the ability to identify and characterize text that satisfies certain kinds of information needs” (Wilbur et al., 2006). The texts in scientific communication in scientific articles consist of sentences constructed to convey information to facilitate deeper understanding of the scientific topic or research.

Sentences in scientific text can be annotated or labelled based on the presence or absence of features such as specific words or group of words defined by the annotator. For example, sentences biomedical text that mention gene or protein symbols could infer molecular relationships between the gene or proteins (Kim et al., 2008, Lee et al., 2013).

2.19 Visual Analytics Facilitated Literature Curation of Arsenic Concentrations in Foods

There is the need to provide computational resources to assist in the manual curation of scientific publications that describe arsenic concentrations in foods. The diverse possibilities for automatic analysis, visualization and interaction with data that is available through visual analytics tools provides a rationale to investigate visual analytics assisted manual curation of text on arsenic in foods. This thesis use the arsenic sentence database, which facilitates the query of the sentences with keywords as well as retrieval of sentences for specific PubMed abstracts (Isokpehi et al., 2010). The resource is available at http://genomics.jsums.edu/sentence/arsenic_pubmed

Visual analytics tools include features to design dashboards (Al-Hajj et al., 2013), which among other functions facilitate decision making through the integration of multiple views as well as external websites (Murray, 2013). An example of a dashboard generated using Tableau is shown in Figure 18.

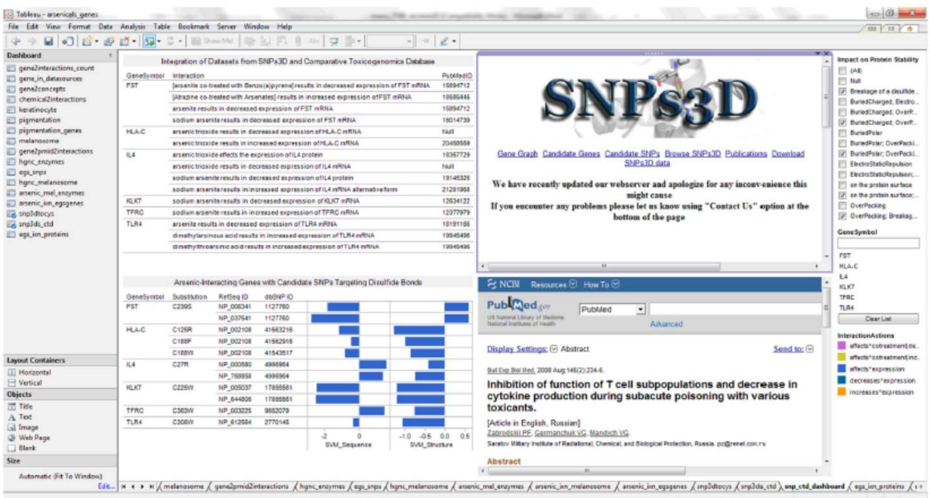


Figure 18. Dashboard illustrating integration of multiple views and external websites.

Source: (Isokpehi et al., 2012)

CHAPTER 3

RESEARCH METHODS

The overall research method is secondary data analyses using visual analytics techniques. Secondary dataset analyses is an established methodology and economical alternative to expensive and time-consuming new data collection projects (Nelson et al., 2013, Pietrobon et al., 2004, Smith et al., 2011). Secondary data analyses can be utilized to investigate additional research questions; inform future research or policy formulation; or develop new analytical approaches secondary to a project's originally intended purpose (NEI, 2012). The possibility of integrating diverse types of data (qualitative and quantitative) using visual analytics provides an opportunity to explore, mine and analyse existing datasets on arsenic in foods.

According to Thomas and Cook (2006) citing Card et al. (1999) and Spence (2000), visual analytics is an analysis dialogue in which the analyst observes the current data representation interprets and makes sense of what he or she sees, and then thinks of the next question to ask, essentially formulating a strategy for how to proceed.

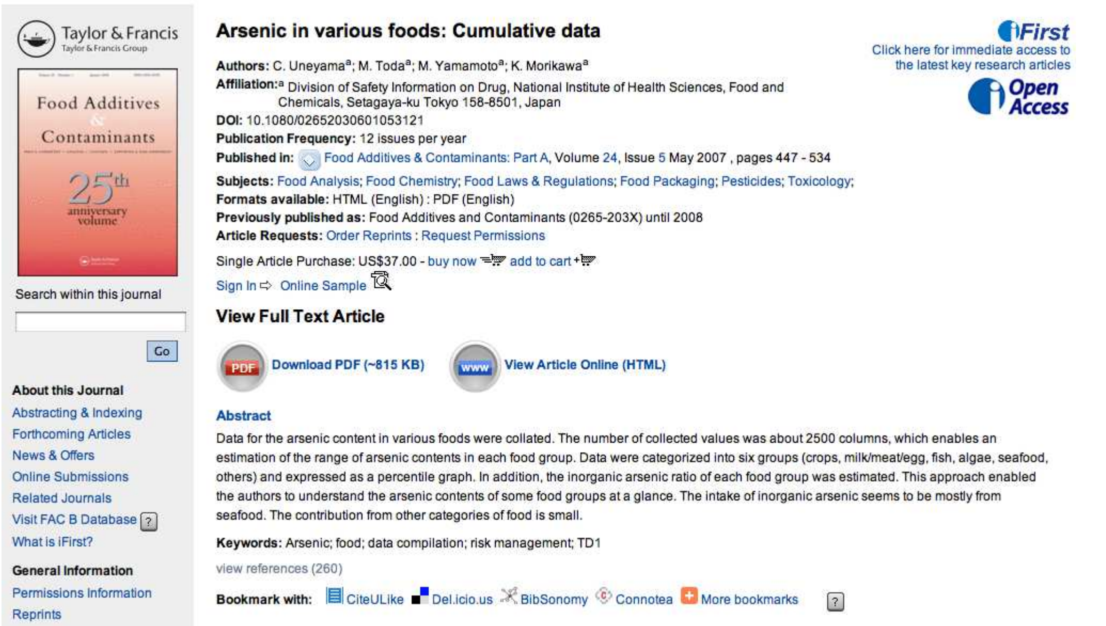
There are a variety of Visual Analytics tools that have been developed to perform visual analytics tasks: (i) analytical reasoning; (ii) visual representations and interaction; (iii) data representations and transformations; and (iv) the production, presentation and dissemination of results of analysis.

Four objectives were investigated in this doctoral research: 1. Provide insightful visual analytic views of compiled data on arsenic in food categories; 2. Categorize table ready foods by arsenic content; 3. Compare arsenic content in rice product categories.

3.1 Objective 1: Provide insightful visual analytic views of compiled data on arsenic in food categories.

3.1.1 Data Collection and Preparation for Visual Analytics

The cumulative data on arsenic in foods collated by Uneyama *et al*. (2007) was the primary data source for visual analytics (Figure 19 and Figure 20). The number of collected values was approximately 2500 rows, which enables an estimation of the range of arsenic contents in each food group. A total of six tables for six food groups (crops, milk/meat/egg, fish, algae, seafood and others) have the following information: (i) name of sample; (ii) country or area for sampling or analysis; (iii) reported or analysed year; (iv) molecular species (if indicated); (v) detection method (if indicated); (vi) number of samples; (vii) original reported values; (vii) conversion for mg kg^{-1} unit; (ix) references, and (x) other/notes. The data in the tables were converted into spreadsheet files in data formats for processing with visual analytics tools.


Taylor & Francis
Taylor & Francis Group
Food Additives & Contaminants
25th anniversary volume
Search within this journal
Go
About this Journal
Abstracting & Indexing
Forthcoming Articles
News & Offers
Online Submissions
Related Journals
Visit FAC B Database
What is iFirst?
General Information
Permissions Information
Reprints

Arsenic in various foods: Cumulative data

iFirst
Click here for immediate access to the latest key research articles
Open Access

Authors: C. Uneyama[a]; M. Toda[a]; M. Yamamoto[a]; K. Morikawa[a]
Affiliation:[a] Division of Safety Information on Drug, National Institute of Health Sciences, Food and Chemicals, Setagaya-ku Tokyo 158-8501, Japan
DOI: 10.1080/02652030601053121
Publication Frequency: 12 issues per year
Published in: Food Additives & Contaminants: Part A, Volume 24, Issue 5 May 2007 , pages 447 - 534
Subjects: Food Analysis; Food Chemistry; Food Laws & Regulations; Food Packaging; Pesticides; Toxicology;
Formats available: HTML (English) : PDF (English)
Previously published as: Food Additives and Contaminants (0265-203X) until 2008
Article Requests: Order Reprints : Request Permissions
Single Article Purchase: US$37.00 - buy now add to cart
Sign In Online Sample

View Full Text Article

PDF Download PDF (~815 KB) www View Article Online (HTML)

Abstract

Data for the arsenic content in various foods were collated. The number of collected values was about 2500 columns, which enables an estimation of the range of arsenic contents in each food group. Data were categorized into six groups (crops, milk/meat/egg, fish, algae, seafood, others) and expressed as a percentile graph. In addition, the inorganic arsenic ratio of each food group was estimated. This approach enabled the authors to understand the arsenic contents of some food groups at a glance. The intake of inorganic arsenic seems to be mostly from seafood. The contribution from other categories of food is small.

Keywords: Arsenic; food; data compilation; risk management; TD1
view references (260)
Bookmark with: CiteULike Del.icio.us BibSonomy Connotea More bookmarks

Figure 19. Screenshot of journal article arsenic in various foods on publisher's website.

Table I. Reported values of for the arsenic content of various crops.

Species	Area	Year	Molecular species	Method	Sample number	Content*	ppm	Reference	Notes
Wheat	USA[†]	1970s	Total arsenic			0.007–0.3 $\mu g\,g^{-1}$	0.007–0.3	Nishizawa (1991b)	
Corn	USA[†]	1970s	Total arsenic			0.001–0.4	0.001–0.4	Nishizawa (1991b)	
Popcorn	USA[†]	1970s	Total arsenic			<0.1	<0.1	Nishizawa (1991b)	
Barley	USA[†]	1970s	Total arsenic			14.3 (wet)	14.3 (wet)	Nishizawa (1991b)	
Rye	USA[†]	1970s	Total arsenic			<0.1	<0.1	Nishizawa (1991b)	
Rice	USA[†]	1970s	Total arsenic			<0.07–3.53	<0.07–3.53	Nishizawa (1991b)	
Rice	USA[†]	1970s	Total arsenic			0.8–5.0	0.8–5.0	Nishizawa (1991b)	
Bread	USA[†]	1970s	Total arsenic			0.016–0.03	0.016–0.03	Nishizawa (1991b)	
Soy	USA[†]	1970s	Total arsenic			0.05–1.22	0.05–1.22	Nishizawa (1991b)	
Soy (for feed)	USA[†]	1970s	Total arsenic			0.07–2.12	0.07–2.12	Nishizawa (1991b)	
Soybean oil	USA[†]	1970s	Total arsenic			0.09	0.09	Nishizawa (1991b)	
Cherry bean (fruit)	USA[†]	1970s	Total arsenic			0.01–0.1	0.01–0.1	Nishizawa (1991b)	
Cherry bean (plant)	USA[†]	1970s	Total arsenic			0.21–0.29	0.21–0.29	Nishizawa (1991b)	
Pea (fruit)	USA[†]	1970s	Total arsenic			<0.01–0.49	<0.01–0.49	Nishizawa (1991b)	
Pea (plant)	USA[†]	1970s	Total arsenic			0.05–22.7	0.05–22.7	Nishizawa (1991b)	
Peanut	USA[†]	1970s	Total arsenic			0.01–0.30	0.01–0.30	Nishizawa (1991b)	
Carrot	USA[†]	1970s	Total arsenic			0.32	0.32	Nishizawa (1991b)	
Potato	USA[†]	1970s	Total arsenic			0.01–0.2	0.01–0.2	Nishizawa (1991b)	
Onion	USA[†]	1970s	Total arsenic			0.015–1.54	0.015–1.54	Nishizawa (1991b)	
Turnip	USA[†]	1970s	Total arsenic			0.036–0.83	0.036–0.83	Nishizawa (1991b)	
Swedish turnip	USA[†]	1970s	Total arsenic			0.8	0.8	Nishizawa (1991b)	
Radish	USA[†]	1970s	Total arsenic			0.01–2.01	0.01–2.01	Nishizawa (1991b)	
Beet	USA[†]	1970s	Total arsenic			0.34	0.34	Nishizawa (1991b)	
Tomato	USA[†]	1970s	Total arsenic			0.01–2.95	0.01–2.95	Nishizawa (1991b)	
Tomato (plant)	USA[†]	1970s	Total arsenic			<0.2–6.75	<0.2–6.75	Nishizawa (1991b)	
Egg plant	USA[†]	1970s	Total arsenic			0.18–0.77	0.18–0.77	Nishizawa (1991b)	
Cucumber	USA[†]	1970s	Total arsenic			0.02–2.4	0.02–2.4	Nishizawa (1991b)	
Lettuce	USA[†]	1970s	Total arsenic			0.14	0.14	Nishizawa (1991b)	
Parsley	USA[†]	1970s	Total arsenic			0.1–8.0	0.1–8.0	Nishizawa (1991b)	
Spinach	USA[†]	1970s	Total arsenic			1.84–2.10	1.84–2.10	Nishizawa (1991b)	
Savoy	USA[†]	1970s	Total arsenic			0.39	0.39	Nishizawa (1991b)	
Cabbage	USA[†]	1970s	Total arsenic			0.0–2.01	0.0–2.01	Nishizawa (1991b)	
Pepper	USA[†]	1970s	Total arsenic			0.39	0.39	Nishizawa (1991b)	
Pepper root	USA[†]	1970s	Total arsenic			1.57	1.57	Nishizawa (1991b)	
Apple	USA[†]	1970s	Total arsenic			0.04–1.72	0.04–1.72	Nishizawa (1991b)	
Orange	USA[†]	1970s	Total arsenic			0.11–0.35	0.11–0.35	Nishizawa (1991b)	
Pear	USA[†]	1970s	Total arsenic			0.17–0.39	0.17–0.39	Nishizawa (1991b)	
Apricot	USA[†]	1970s	Total arsenic			0.15–1.5	0.15–1.5	Nishizawa (1991b)	
Lemon	USA[†]	1970s	Total arsenic			0.5	0.5	Nishizawa (1991b)	

(continued)

Arsenic in various foods: Cumulative data 449

Figure 20. Example of dataset on arsenic in foods.

Source: (Uneyama et al., 2007).

3.1.2 Use Cases for Visual Analytics of Datasets

Use cases for interactive analysis of the dataset with emphasis on rice were developed from the food group compilation. The use cases that were developed are (i) identify crops that were tested for inorganic arsenic; (ii) group rice samples by molecular species of arsenic; and (iii) group rice samples by country of origin. For each use case, the data field use as filter is described, followed purpose of the use cases and justification from prior research. In three use cases, extensive filtering, grouping and rearrangement of data fields (columns) were performed. This visual representation and interaction tasks could not be easily performed in a spreadsheet software.

In the first use case, the data field used as filter was molecular species, which provides arsenic species (such as arsenite, arsenate, organic arsenic, DMA, MMA and total arsenic) that were tested for each crop sample. This use case will allow the evaluation of the inorganic arsenic content in multiple crops. Consumer advocacy groups increasing seek to know the type of arsenic species present in the food or food source (Jackson et al., 2012). In humans, inorganic arsenic interact with cellular process thereby affecting the normal function of cells (Hughes, 2006).

In the second use case, the data field used as filter was the crop (species in the dataset). The availability of diverse rice varieties provides a need for easy comparison of the arsenic content of rice varieties. The use case will allow for comparison of the arsenic content (ppm) of different types of rice. Rice has more arsenic content compared to other grains (Sun et al., 2008, Wei et al., 2013). In the third use case, the data field used as filter was Area. The Area could also mean the country of origin of the crop species tested for arsenic content. In addition to comparison of arsenic contents, outlier values from other data fields could be identified. There is variation in the content of arsenic in rice depending on where the rice is cultivated (Meharg et al., 2009).

3.2. Objective 2: Categorize table ready foods by arsenic content.

3.2.1 Data Collection and Preparation for Visual Analytics

Two types of TDS data sources were collected and processed with visual analytics tasks. These data sources and processing to format suitable for visual analytics tasks are described below.

United States Food and Drug Administration Total Diet Study (1991-3 through 2005-4. The statistics on the toxic and nutritional elements found in the United States TDS foods are available as Portable Document Format (PDF) files, providing opportunities for data integration, visualization and analysis (Figure 21). The summary of results for arsenic content in the TDS (March 1991 to April 2005) consisted of 328 records and 10 data fields. The data fields (table columns) were TDS Food Description, TDS Food No., Number of Results, Number Not Detected, Number of Traces, Mean (mg/kg), Std Dev [Standard Deviation] (mg/kg), Minimum (mg/kg), Maximum (mg/kg) and Median (mg/kg). Each data record (row) contains data for each of the data fields. For the visual analytics, two datasets were constructed as spreadsheet files in Microsoft Excel (Microsoft Corporation Richmond, WA, USA). The first data set named "element_tds" consisted of all the data fields except TDS Food Description. The second dataset named "tds_foods" consisted of data fields: TDS Food Description, TDS Food No. and Food Category. Each TDS Description and its TDS Food No. were mapped to a Food Category by the author using guidelines provided by the U.S. FDA (Egan et al., 2007). The categories were Dairy products; Eggs; Baby foods; Meat, poultry, fish; Legumes; Grains; Fruits; Vegetables; Mixtures; Sweets; Fats/oils; and Beverages.

US Food and Drug Administration — Total Diet Study — Market Baskets 1991-3 through 2005-4

Arsenic - Summary of Results

TDS Food Description	TDS Food No.	Number of Results	Number Not Detected	Number of Traces	Mean (mg/kg)	Std Dev (mg/kg)	Minimum (mg/kg)	Maximum (mg/kg)	Median (mg/kg)
Overall:		13,231	11,663	969	0.024	0.250	0	10.4	0
whole milk, fluid	1	51	51	0	0	0	0	0	0
lowfat (2% fat) milk, fluid	2	51	51	0	0	0	0	0	0
chocolate milk, fluid	3	51	51	0	0	0	0	0	0
skim milk, fluid	4	51	49	2	0.001	0.003	0	0.020	0
plain yogurt, lowfat	6	39	39	0	0	0	0	0	0
chocolate milk shake, fast-food	7	51	51	0	0	0	0	0	0
evaporated milk, canned	8	39	38	1	0	0.002	0	0.012	0
American, processed cheese	10	51	51	0	0	0	0	0	0
cottage cheese, 4% milkfat	11	39	39	0	0	0	0	0	0
cheddar cheese	12	51	51	0	0	0	0	0	0
ground beef, pan-cooked	13	51	48	3	0.001	0.003	0	0.017	0
beef chuck roast, baked	14	51	48	3	0.001	0.003	0	0.012	0
beef steak, loin, pan-cooked	16	39	34	5	0.002	0.005	0	0.018	0
ham, baked	17	51	50	1	0	0.002	0	0.011	0
pork chop, pan-cooked	18	51	51	0	0	0	0	0	0
pork sausage, pan-cooked	19	51	50	0	0.001	0.008	0	0.055	0
pork bacon, pan-cooked	20	51	51	0	0	0	0	0	0
pork roast, baked	21	51	51	0	0	0	0	0	0
lamb chop, pan-cooked	22	51	51	0	0	0	0	0	0
chicken, fried (breast, leg, and thigh) homemade	24	39	12	21	0.020	0.021	0	0.086	0.017
turkey breast, roasted	26	51	26	22	0.011	0.016	0	0.078	0
liver, beef, fried	27	51	37	13	0.005	0.011	0	0.055	0
frankfurters, beef, boiled	28	51	51	0	0	0	0	0	0
bologna, sliced	29	51	51	0	0	0	0	0	0
salami, sliced	30	51	51	0	0	0	0	0	0
tuna, canned in oil	32	39	1	0	0.929	0.326	0	1.71	0.910
fish sticks, frozen, heated	34	51	0	0	0.736	0.475	0.130	2.79	0.674

Page 3 of 211 Rev 4.1

Figure 21. Section of the data on concentration of arsenic in table ready foods.

Source: http://www.fda.gov/downloads/Food/FoodScienceResearch/TotalDietStudy/UCM243059.pdf

Detailed food by food data for arsenic intake by total diet study. In the review article on arsenic in various foods (Uneyema et al. 2007), Tables X to XIII included estimates of inorganic arsenic intake from foods in the TDS conducted in several countries. Figure 22 is an example of a table on detailed food by food data for arsenic intake from a 1990 TDS conducted in Spain. Four datasets (Table X, United Kingdom, 1997; Table XI, Canada 1985-1988; Table XII, Spain, 1990; and Table XII, Okinawa, Japan, 2000) were identified for processing. Data from two data fields were obtained from each table and integrated into one data set in the spreadsheet file. The data fields for the data set named “uneyema_tds” were Food group, Country, Year and Estimated Daily inorganic Arsenic Intake (µg).

Table XII. Detailed food by food data for total arsenic intake by total diet study: Spain, 1990.

Food group	Average intake (g)	As concentration ($\mu g\,kg^{-1}$)	As daily intake (μg)	Ratio (%)	Estimated daily inorganic As intake (μg)	Ratio (%)
Eggs	41	<5	0	0.00	0.00	0.00
Meat	118	<5	1	0.39	1.00	6.32
Meat products	45	<5	0	0.00	0.00	0.00
Fish	89	2800	249	97.65	10.46	66.11
Milk	294	<2	1	0.39	1.00	6.32
Dairy products	58	<5	0	0.00	0.00	0.00
Bread	122	<5	1	0.39	0.84	5.31
Cereals	62	<5	0	0.00	0.00	0.00
Pulses and nuts	27	<5	0	0.00	0.00	0.00
Potatoes	90	<3	1	0.39	0.84	5.31
Vegetables	159	<3	1	0.39	0.84	5.31
Fruits	377	<2	1	0.39	0.84	5.31
Sugar and preserves	34	<5	0	0.00	0.00	0.00
Fats and oils	45	<5	0	0.00	0.00	0.00
Non-alcoholic beverages	198	<1	0	0.00	0.00	0.00
Alcoholic beverages	243	2	0	0.00	0.00	0.00
Total			255	100.00	15.818	100.00

Detailed food by food data for total arsenic intake by total diet study of Spain. The calculation method is same as in Table X. Data are energy content 10 390 kJ day^{-1}. Source: Urieta et al. (1996).

Figure 22. Example of detailed food by food data for arsenic intake by total diet study.

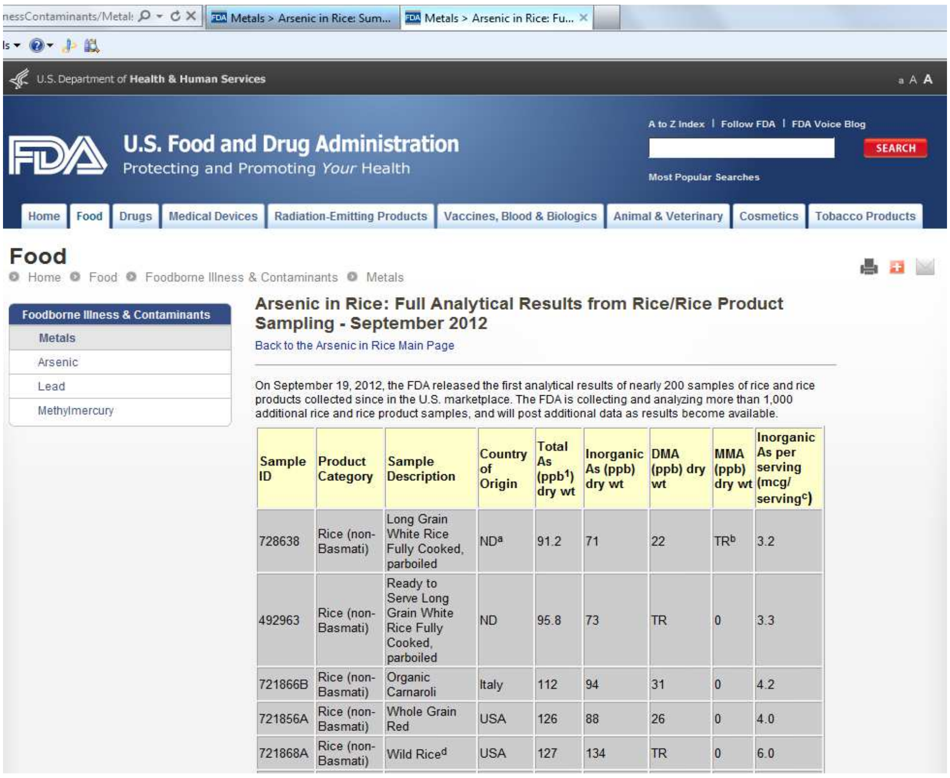
3.2.2. Use Cases for Visual Analytics of Datasets

Three use cases for visual analytics were developed for the U.S. FDA TDS dataset. These are (i) identify foods with arsenic; (ii) identify foods without arsenic; and (iii) group foods by category and arsenic content. The use case for the Uneyema TDS dataset was to compare the estimated daily inorganic arsenic intake (µg) for food group from countries.

3.3. Objective 3: Compare arsenic content in rice product categories.

3.3.1 Data Collection and Preparation for Visual Analytics

On September 19, 2012, the U.S. FDA released the first analytical results of arsenic (As) content of nearly 200 samples of rice and rice products collected in the U.S. marketplace. The results are available on the U.S. FDA website (www.fda.gov) (Figure 23). The dataset of 193 records consisted of the following data fields: Sample ID, Product Category; Sample Description; Country of Origin; Total As (ppb) dry wt; Inorganic As (ppb) dry wt; DMA (ppb) dry wt; MMA (ppb) dry wt; and Inorganic As per serving (mcg/serving). The dataset was obtained from the website and copied to word processing software and then to a spreadsheet file where all superscripts to notes were removed. The dataset of 193 columns and 9 fields were stored as a spreadsheet file for visual analytics.



Arsenic in Rice: Full Analytical Results from Rice/Rice Product Sampling - September 2012

Back to the Arsenic in Rice Main Page

On September 19, 2012, the FDA released the first analytical results of nearly 200 samples of rice and rice products collected since in the U.S. marketplace. The FDA is collecting and analyzing more than 1,000 additional rice and rice product samples, and will post additional data as results become available.

Sample ID	Product Category	Sample Description	Country of Origin	Total As (ppb[1]) dry wt	Inorganic As (ppb) dry wt	DMA (ppb) dry wt	MMA (ppb) dry wt	Inorganic As per serving (mcg/ serving[c])
728638	Rice (non-Basmati)	Long Grain White Rice Fully Cooked, parboiled	ND[a]	91.2	71	22	TR[b]	3.2
492963	Rice (non-Basmati)	Ready to Serve Long Grain White Rice Fully Cooked, parboiled	ND	95.8	73	TR	0	3.3
721866B	Rice (non-Basmati)	Organic Carnaroli	Italy	112	94	31	0	4.2
721856A	Rice (non-Basmati)	Whole Grain Red	USA	126	88	26	0	4.0
721868A	Rice (non-Basmati)	Wild Rice[d]	USA	127	134	TR	0	6.0

Figure 23. Section of the data on concentration of arsenic in rice samples collected in the United States marketplace.

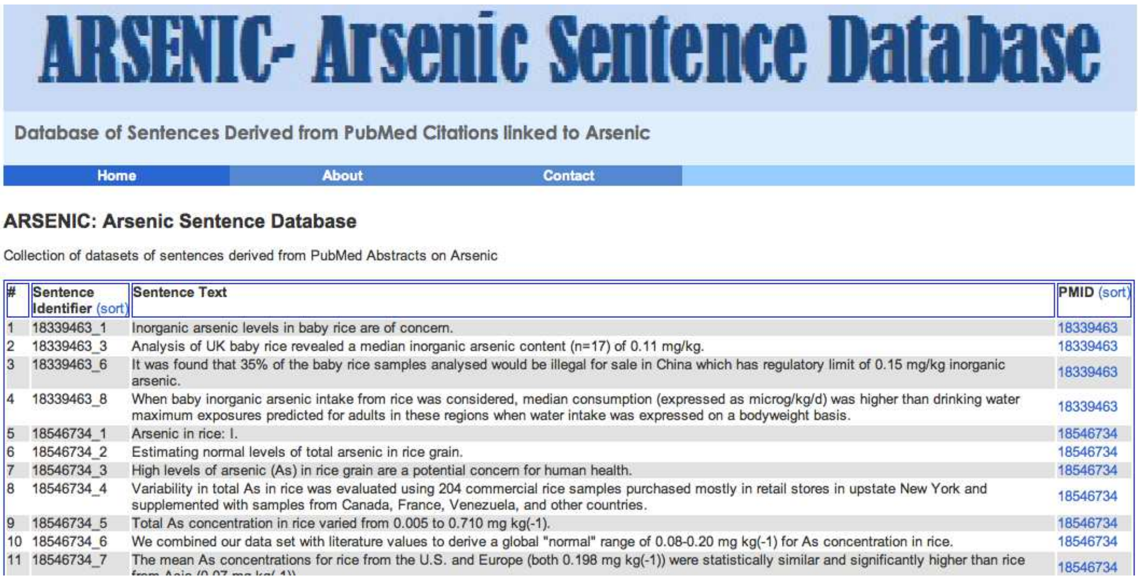
3.2.2 Use Cases for Visual Analytics of Datasets

The use case for visual analytics of the US FDA Arsenic in Rice was to compare rice product categories for Inorganic Arsenic per serving. Views were constructed to display the inorganic arsenic content of the rice samples.

3.4. Objective 4: Identify informative sentences on arsenic concentrations in rice.

3.4.1 Data Collection and Preparation for Visual Analytics

Search of the Arsenic Sentence Database (Isokpehi et al., 2010) (http://genomics.jsums.edu/sentence/arsenic_pubmed) for sentences with keyword "rice" retrieved 758 sentences from 240 PubMed abstracts. Figure 24 shows a screenshot of the Arsenic Sentence Database with result of search with "rice" as keyword. The assumption of this research is that the PubMed abstracts and associated full-text articles will contain informative text fragments on arsenic concentrations in rice. The focus of this objective was to identify sentence that include concentration of arsenic.



#	Sentence Identifier (sort)	Sentence Text	PMID (sort)
1	18339463_1	Inorganic arsenic levels in baby rice are of concern.	18339463
2	18339463_3	Analysis of UK baby rice revealed a median inorganic arsenic content (n=17) of 0.11 mg/kg.	18339463
3	18339463_6	It was found that 35% of the baby rice samples analysed would be illegal for sale in China which has regulatory limit of 0.15 mg/kg inorganic arsenic.	18339463
4	18339463_8	When baby inorganic arsenic intake from rice was considered, median consumption (expressed as microg/kg/d) was higher than drinking water maximum exposures predicted for adults in these regions when water intake was expressed on a bodyweight basis.	18339463
5	18546734_1	Arsenic in rice: I.	18546734
6	18546734_2	Estimating normal levels of total arsenic in rice grain.	18546734
7	18546734_3	High levels of arsenic (As) in rice grain are a potential concern for human health.	18546734
8	18546734_4	Variability in total As in rice was evaluated using 204 commercial rice samples purchased mostly in retail stores in upstate New York and supplemented with samples from Canada, France, Venezuela, and other countries.	18546734
9	18546734_5	Total As concentration in rice varied from 0.005 to 0.710 mg kg(-1).	18546734
10	18546734_6	We combined our data set with literature values to derive a global "normal" range of 0.08-0.20 mg kg(-1) for As concentration in rice.	18546734
11	18546734_7	The mean As concentrations for rice from the U.S. and Europe (both 0.198 mg kg(-1)) were statistically similar and significantly higher than rice	18546734

Figure 24. Screenshot of section of results from Arsenic Sentence Database using rice as search term.

3.4.2 Identification of Sentences with Information on Arsenic Concentration

The 758 sentences from the Arsenic Sentence Database was obtained from (http://genomics.jsums.edu/sentence/arsenic_pubmed) (Isokpehi et al., 2010). The set of sentences uploaded to a spreadsheet software for annotation as informative (Y) or non-informative (N) for arsenic concentration in rice. Thus Y or N was entered in a column named informative for content. The annotation was done by the author. The spreadsheet file with the annotated sentences was then uploaded into Tableau Software for visual analytics tasks.

CHAPTER 4

RESULTS

4.1 Objective 1: Provide insightful visual analytic views of compiled data on arsenic in foods.

4.1.1 Dataset Description

The data table on arsenic in crops provided by Uneyema et al. (2007) was the focus of this objective. The dataset consisted of 459 rows in 9 columns. The fields in the table were Area, Content, Method, Molecular Species, ppm (parts per million), References, Sample Number, Species and Year. Additional information on the description, types and numbers of records associated with each dimension is presented in Table 8.

Table 8. Dataset characteristics for arsenic in crops

Dimension	Description	Number of data types in Dimension
Area	Origin of Food	37
Content	Arsenic content reported in article	356
Method	Detection method of arsenic	17
Molecular Species	Arsenic species	7
Ppm	Arsenic content in Parts per Million	351
References	Journal Article Source of Information	60
Sample Number	Number of Samples Analyzed	44
Species	Crops	279
Year	Year(s) in which Study was Performed	32

4.1.2 Use Case: Identify crops that were tested for inorganic arsenic

The molecular species recorded and the number of records in the dataset were Arsenic (1), As(III) (1), As(V) (1), DMA (9), Inorganic arsenic (32), MMA (1) and Total arsenic (414). The design for this use case consisted of arranging the dimensions in the following order: Area, Year, References, Method, Species and ppm (Figure 25). The number of times the same record is encountered in the dataset is also included (Figure 26).

The visualization revealed three study locations: Chile, Mexico and USA that measured the presence of inorganic arsenic in foods. In the report from Chile, 10 crops were retrieved with the arsenic content (in ppm) ranged from 0.013 (Cabbage) to 0.613 (Spinach). Alfalfa leaf and Alfalfa root were analysed in the study in Mexico with arsenic content ranging from 0.5 to 2 ppm and 1 to 3 ppm respectively. The two reports from the USA were conducted in 2001 and 2004 on rice varieties (long grain white, long grain brown, wild rice, instant long grain, rice flour and cooked rice). The range of arsenic contents reported for the rice samples were 0.0231 to 0.108 ppm.

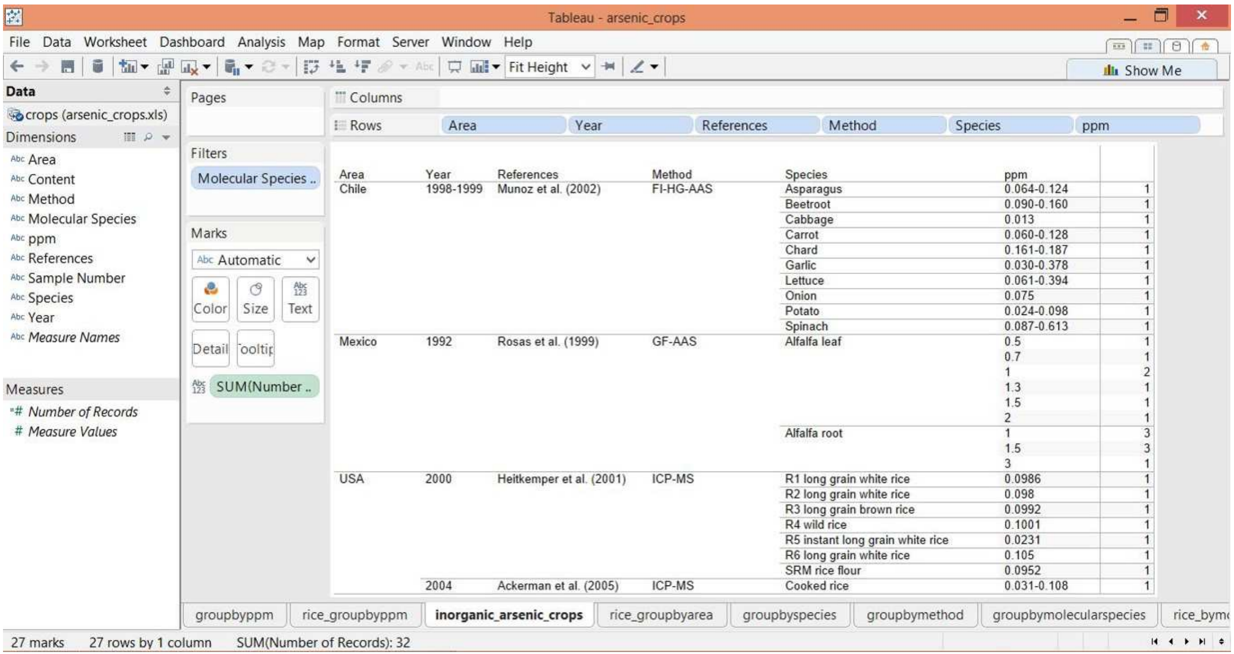


Figure 25. Arrange of dimensions for insights on inorganic arsenic in crops

Area	Year	References	Method	Species	ppm	
Chile	1998-1999	Munoz et al. (2002)	FI-HG-AAS	Asparagus	0.064-0.124	1
				Beetroot	0.090-0.160	1
				Cabbage	0.013	1
				Carrot	0.060-0.128	1
				Chard	0.161-0.187	1
				Garlic	0.030-0.378	1
				Lettuce	0.061-0.394	1
				Onion	0.075	1
				Potato	0.024-0.098	1
				Spinach	0.087-0.613	1
Mexico	1992	Rosas et al. (1999)	GF-AAS	Alfalfa leaf	0.5	1
					0.7	1
					1	2
					1.3	1
					1.5	1
					2	1
				Alfalfa root	1	3
					1.5	3
					3	1
USA	2000	Heitkemper et al. (2001)	ICP-MS	R1 long grain white rice	0.0986	1
				R2 long grain white rice	0.098	1
				R3 long grain brown rice	0.0992	1
				R4 wild rice	0.1001	1
				R5 instant long grain white rice	0.0231	1
				R6 long grain white rice	0.105	1
				SRM rice flour	0.0952	1
	2004	Ackerman et al. (2005)	ICP-MS	Cooked rice	0.031-0.108	1

Figure 26. View showing inorganic arsenic concentration in selected crops.

Notes: Dimensions are arranged in order group the crops according to Area, Year, References, Method, Species and ppm. The numbering at the end is the number of times the record with identical data occurred. An insight is Areas where investigation on inorganic arsenic crops have been conduct. http://public.tableausoftware.com/views/mj_arsenic_crops/inorganic_arsenic_crops

4.1.3 Use Case: Group rice samples by molecular species of arsenic

The design of the view consisted of arrangement of the dimensions in row in the following order: Molecular Species, Method, References, Area, ppm and Species. Filters were provided for the all the dimensions. To obtain only records that are associated with rice samples, the word rice was included in the search box for the filter for Molecular Species. The view shows the concentration of the molecular species of arsenic for 7 rice samples: R1 long grain white rice; R2 long grain white rice; R3 long grain brown rice; R4 wild rice; R5 instant long grain white rice; R6 long grain white rice; and SRM rice flour (Figure 27).

The dataset can be grouped into 7 groups based on the molecular species of arsenic analyzed in the studies. One report (Heitkemper et al. 2001) was identified to have evaluated three molecular species of arsenic (Dimethylarsinic acid [DMA]), inorganic arsenic and Total arsenic) providing an opportunity to compare the content of these three molecular species in rice samples analysed.

Therefore a new view was constructed where the dimensions were arranged on the row of the Tableau Workbook in the following order: Molecular Species, Method, References, Area, ppm and Species. The filters for all the fields were also included. The word "rice" was used as filter for Species dimension while the report Heitkemper et al. (2001) was the filter for Reference dimension.

Molecular Species	Method	References	Area	ppm	Species	
DMA	ICP-MS	Heitkemper et al. (2001)	USA	0.182	R2 long grain white rice	1
				0.195	R6 long grain white rice	1
				0.0099	R4 wild rice	1
				0.0608	R3 long grain brown rice	1
				0.1764	SRM rice flour	1
				0.1869	R5 instant long grain white rice	1
				0.2414	R1 long grain white rice	1
Inorganic arsenic	ICP-MS	Heitkemper et al. (2001)	USA	0.098	R2 long grain white rice	1
				0.105	R6 long grain white rice	1
				0.0231	R5 instant long grain white rice	1
				0.0952	SRM rice flour	1
				0.0986	R1 long grain white rice	1
				0.0992	R3 long grain brown rice	1
				0.1001	R4 wild rice	1
Total arsenic	ICP-MS	Heitkemper et al. (2001)	USA	0.3+-0.01	R6 long grain white rice	1
				0.11+-0.01	R4 wild rice	1
				0.16+-0.01	R3 long grain brown rice	1
				0.21+-0.01	R5 instant long grain white rice	1
				0.28+-0.01	R2 long grain white rice	1
				0.28+-0.02	SRM rice flour	1
				0.34+-0.02	R1 long grain white rice	1

Figure 27. Comparison of content of arsenic species in rice varieties.

Notes: An insight is the noteworthy difference in the concentration of DMA in wild rice compared to other conventional rice varieties.
http://public.tableausoftware.com/views/mj_arsenic_crops/molspec_compare

In all the six rice grains analysed, the total arsenic content in ppm was higher for long grain white rice samples (0.21+-0.01 to 0.34+-0.02) compared with brown long grain rice (0.16+-0.01) and wild rice (0.11+-0.01). An outlier concentration of 0.0099 ppm was observed for DMA concentration in R4 wild rice.

4.1.5. Use Case: Group rice samples by country of origin.

The view designed arranged the fields as follows: Species, ppm, Area, Method, References and Year. Only records for Total Arsenic and “Rice” were displayed. A total of 23 records were displayed. To facilitate comparison of the arsenic concentrations, the records were exported to a spreadsheet file for data entry. The maximum total arsenic concentration was identified from each record under a field (maximum ppm). The spreadsheet file was uploaded as a data source into Tableau Software. Additional comparisons were the performed on the data source. The view designed arranged in the following order: Area, References, Year, Method and Concentration. References s presented in Figure 28. Records without a maximum value or references not in English Language were excluded for the visualized data. The view revealed an investigation in Malaysia where the total arsenic concentration in rice was 2.590 ppm, the highest in the dataset. This outlier observation provides basis to discuss latest data on arsenic in rice locally produced Malaysia.

Area	References	Year	Method	Maximum ppm
Bangladesh	Bae et al. (2002)	2001	HG-AAS	0.173
	Das et al. (2004)	2003	HG-AAS	0.140
	Duxbury et al. (2003)	2000	ICP-MS	0.420
		2002	HG-ICP	0.420
	Hironaka and Ahmad (2000)	1998-1999	AAS	0.940
	Meharg and Rahman (2003)	2001	HG-AAS	1.835
	Williams et al. (2005)	2004	ICP-MS	0.300
German	Kohlmeyer et al. (2003)	2003	ICP-MS	0.500
Japan	Hironaka and Ahmad (2000)	1998-1999	AAS	0.320
	Onozuka et al.et(2001)	1998-2001	AAS	0.290
Japan (unpolluted area)	Onozuka et al.et(2001)	1998-2001	AAS	0.250
Malaysia	Zarcinas et al. (2004a)	2001	ICPS	2.590
Spain	Pizarro et al. (2003)	2002	HG-AFS	0.283
Tokyo, Japan	Onozuka et al.et(2000)	1998-2000	AAS	0.290
	Onozuka et al.et(2003)	1999	AAS	0.620
		2000	AAS	0.510
		2001	AAS	0.550
		2002	AAS	0.440
USA	Williams et al. (2005)	2004	ICP-MS	0.400

0.5 1.0 1.5 2.0 2.5

Maximum ppm

Figure 28. Outlier findings for concentrations of rice obtained from several areas (countries)

Notes: An insight is the noteworthy difference in the concentration of DMA in wild rice compared to other conventional rice varieties.
http://public.tableausoftware.com/views/mj_arsenic_crops/ricebyarea_maxppm

4.2 Objective 2: Categorize table ready foods by arsenic content.

4.2.1 Dataset Description

The 328 table ready foods from the US FDA Total Diet Study were categorized into 12 food categories: baby foods (37); beverages (15); dairy (25); eggs (3); fats/oils (12); fruits (35); grains (55); legumes (12); meat, poultry, fish (37); mixtures (35); sweets (12); and vegetables (50) (Figure 29). The count of unique values for the minimum, mean and maximum arsenic content (mg/kg) were 11, 38 and 27 respectively.

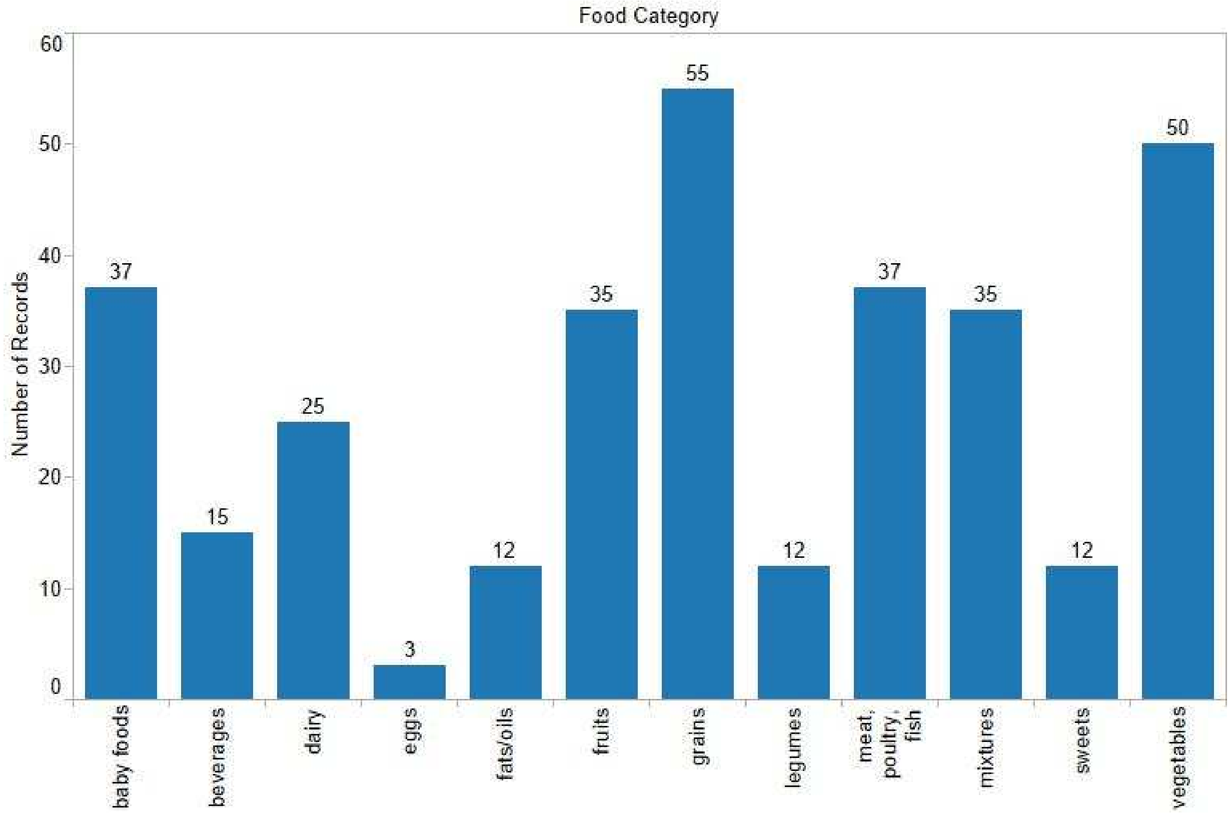


Figure 29. Count of table ready foods in 12 food categories.

4.2.2 Use Case: Identify foods with arsenic

Table ready foods with arsenic content are be those with a minimum arsenic content greater than 0. A total of 185 foods were identified to contain arsenic. Filtering for “rice” in the Total Diet Study (TDS) Food Description resulted in 10 table ready foods in 3 categories (Figure 30): baby foods (4 foods), grains (3 foods) and mixtures (3 foods).

A complete list of foods with detected arsenic concentration is available in Appendix 1. Since rice belongs to the food category of grains, the concentrations of arsenic in grains in the TDS is provided in Figure 31.

Food Category	Maximum (mg/kg)	TDS Food No	TDS Food Description	
baby foods	0.02	319	rice cereal with apple, strained/junior	1
	0.03	312	rice cereal, strained/junior	1
	0.047	216	turkey and rice, strained/junior	1
	0.087	311	rice infant cereal, instant, prepared with whole milk	1
grains	0.054	324	BF, cereal, rice, dry, prep w/ water	1
	0.128	50	white rice, cooked	1
	0.32	75	crisped rice cereal	1
mixtures	0.044	325	BF, cereal, rice w/apples, dry, prep w/ water	1
	0.06	270	green peppers stuffed with beef and rice, homemade	1
	0.106	364	Fried rice, meatless, from Chinese carry-out	1

Figure 30. Inorganic arsenic content of table ready foods prepared with rice.

Notes: View available at http://public.tableausoftware.com/views/mj_usa_tds/with_arsenic

Food Category	Maximum (mg/kg)	TDS Food No	TDS Food Description	
grains	0.011	51	oatmeal, quick (1-3 min), cooked	1
		60	cornbread, homemade	1
	0.012	67	corn chips	1
		68	pancake from mix	1
		149	spaghetti with tomato sauce, canned	1
	0.013	252	butter-type crackers	1
	0.015	347	Spaghetti, enriched, boiled	1
	0.017	64	rye bread	1
	0.019	344	Pancakes, frozen, heated	1
	0.021	59	white roll	1
		63	tortilla, flour	1
		73	shredded wheat cereal	1
		289	chocolate snack cake with chocolate	1
	0.023	66	saltine crackers	1
	0.024	184	sandwich cookies with creme filling, commercial	1
	0.027	291	brownies, commercial	1
	0.028	369	Cake, yellow w/ icing	1
	0.029	294	pretzels, hard, salted, any shape	1
	0.031	250	English muffin, plain, toasted	1
	0.034	62	whole wheat bread	1
	0.037	74	raisin bran cereal	1
	0.039	182	sweet roll/Danish, commercial	1
	0.041	251	graham crackers	1
	0.043	370	Granola bar, w/ raisins	1
	0.045	77	oat ring cereal	1
	0.047	72	fruit-flavored, sweetened cereal	1
	0.052	248	cracked wheat bread	1
	0.054	76	granola cereal	1
		324	BF, cereal, rice, dry, prep w/ water	1
	0.128	50	white rice, cooked	1
	0.286	55	corn, canned	1
	0.32	75	crisped rice cereal	1

Figure 31. Inorganic content of foods categorized as grains.

Notes: An insight is group of foods with identical maximum concentration of inorganic arsenic.
http://public.tableausoftware.com/views/mj_usa_tds/grpbycatg_conc

4.2.3 Use Case: Identify foods without arsenic

A total of 143 table ready foods had a maximum arsenic content of 0. Additionally, there was no food record displayed when the dataset was filtered for “rice” in the Total Diet Study (TDS) Food Description. A subset of table ready foods made from grain that do not contain arsenic is presented in Figure 32. A complete list of foods without detected arsenic concentration is available in Appendix 2.

Food Category	Maximum (mg/kg)	TDS Food No	TDS Food Description	
grains	0	52	wheat cereal, farina, quick (1-3min), cooked	1
		53	corngrits, regular, cooked	1
		54	corn, fresh/frozen, boiled	1
		56	cream style corn, canned	1
		57	popcorn, popped in oil	1
		58	white bread	1
		61	biscuit, from refrigerated dough, baked	1
		65	blueberry muffin, commercial	1
		69	egg noodles, boiled	1
		70	macaroni, boiled	1
		71	corn flakes	1
		178	chocolate cake with chocolate icing, commercial	1
		179	yellow cake with white icing, prepared from cake and icing..	1
		183	chocolate chip cookies, commercial	1
		249	bagel, plain	1
		290	cake doughnuts with icing, any flavor, from doughnut store	1
		292	sugar cookies, commercial	1
		323	BF, cereal, oatmeal, dry, prep w/ water	1
		345	Breakfast tart/toaster pastry	1
		346	Macaroni salad, from grocery/deli	1
		367	Soup, Oriental noodles (ramen noodles), prep w/ water	1
		372	Popcorn, microwave, butter-flavored	1
		374	Brown gravy, canned or bottled	1

Figure 32. Table ready foods in grain food category without inorganic arsenic

Notes: The foods could be alternative to foods in the same category that contain arsenic.
http://public.tableausoftware.com/views/mj_usa_tds/without_arsenic

4.2.4 Use Case: Group foods by category and arsenic content

The final design of the view arranged the fields as follows: Maximum (mg/kg), Mean (mg/kg), Food Category, TDS Food No, TDS Food Description. This view enabled the grouping of foods that had identical Maximum arsenic concentration. The inclusion of the Mean arsenic concentration provides an additional decision-making feature. A subset of the dataset is presented in Figure 33. The groups include grains and have a maximum arsenic concentration of at least 0.030.

Maximum (mg/kg)	Mean (mg/kg)	Food Category	TDS Food No	TDS Food Description	
0.031	0.001	grains	250	English muffin, plain, toasted	1
		legumes	246	peas, mature, dry, boiled	1
0.034	0.002	baby foods	211	vegetables and beef, strained/junior	1
	0.006	grains	62	whole wheat bread	1
0.037	0.003	fruits	97	avocado, raw	1
	0.011	grains	74	raisin bran cereal	1
	0.021	meat, poultry, fish	336	Chicken breast, fried, fast-food (w/	1
0.039	0.001	grains	182	sweet roll/Danish, commercial	1
	0.01	mixtures	282	beef chow mein, from Chinese carry	1
0.043	0.001	dairy	235	fruit-flavored yogurt, lowfat (fruit	1
		vegetables	259	carrot, fresh, boiled	1
	0.003	vegetables	107	spinach, fresh/frozen, boiled	1
			267	okra, fresh/frozen, boiled	1
	0.027	grains	370	Granola bar, w/ raisins	1
0.047	0.003	baby foods	216	turkey and rice, strained/junior	1
	0.004	mixtures	284	mushroom soup, canned, condensed, prepared with whole milk	1
	0.006	grains	72	fruit-flavored, sweetened cereal	1
0.054	0.022	grains	76	granola cereal	1
	0.046	grains	324	BF, cereal, rice, dry, prep w/ water	1

Figure 33. Groups of foods with maximum inorganic arsenic content of at least 0.03 mg/kg.

Notes: Inclusion of mean inorganic arsenic concentration splits the group.
http://public.tableausoftware.com/views/mj_usa_tds/arsenic_grains

4.2.5 Use Case: Compare the estimated daily inorganic arsenic intake (µg) for food group from countries.

Four tables were obtained from the publication of Uneyema et al. (2007). An integration of data from two fields of the TDS datasets for four countries is presented in Figure 34 (Table X, United Kingdom, 1997; Table XI, Canada 1985-1988; Table XII, Spain, 1990; and Table XII, Okinawa, Japan, 2000). A noteworthy observation for the view is that from Japan TDS in 2000, the daily inorganic arsenic intake for rice was 15.87 µg and the highest for the foods compared. The concentration for rice was 10.46 followed by fish from a 1990 TDS in Spain. In the UK 1997 TDS, fish had the highest estimate of 2.56 µg.

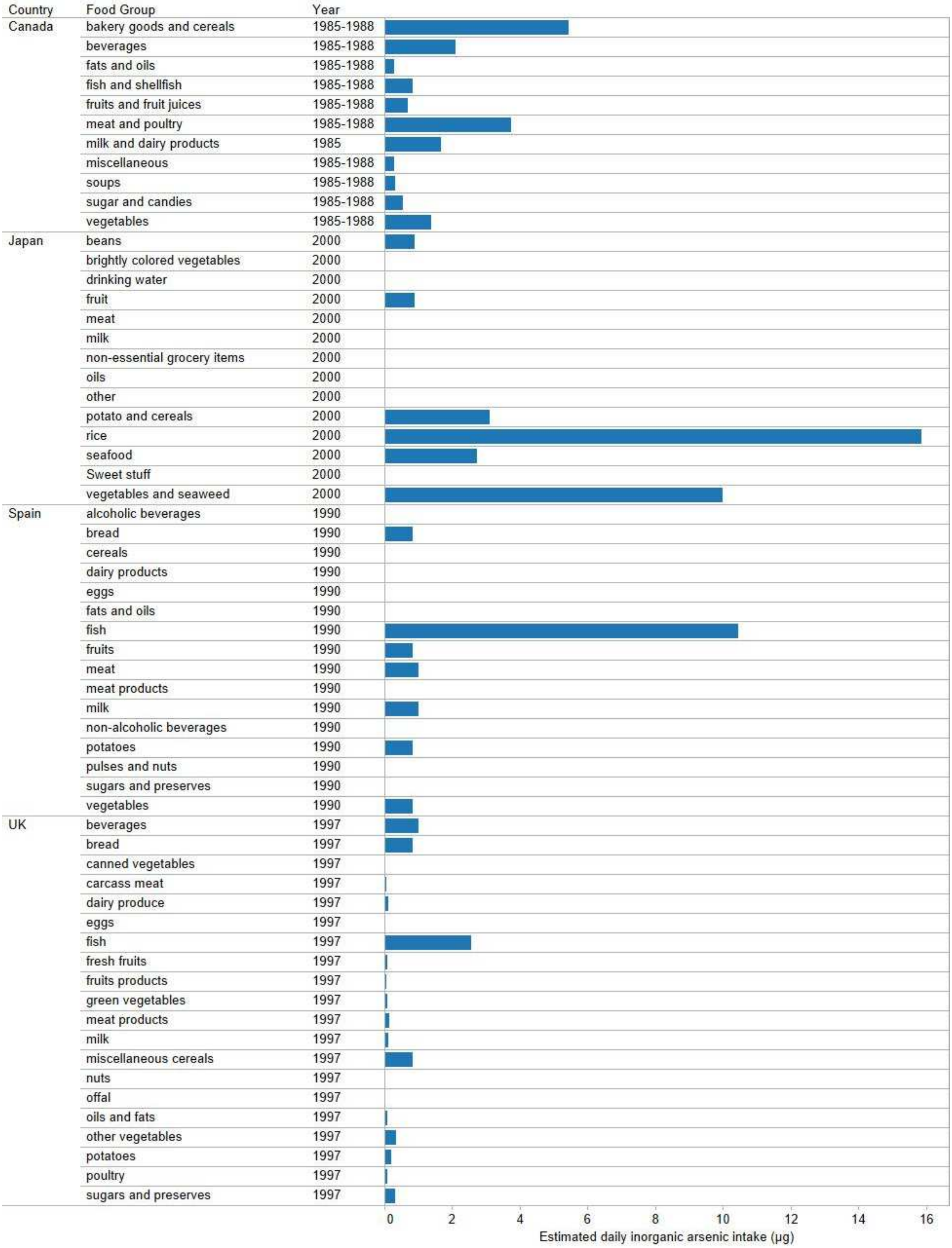


Figure 34. Comparison of total diet studies for estimated daily inorganic arsenic intake.

Notes: A noteworthy observation for the view is that from Japan TDS in 2000, the daily inorganic arsenic intake for rice was 15.87 µg and the highest for the foods compared.
http://public.tableausoftware.com/views/mj_usa_tds/uneyema_tds01

An additional view was constructed to compare the estimated daily inorganic intake for the food groups from the total diet studies (Figure 35). A purpose of this view is to identify food groups reported by multiple countries. In the view, the countries are represented by bars of varying width. The following food groups were included in TDS from more than one country: beverages, bread, fish, milk, potatoes and vegetables.

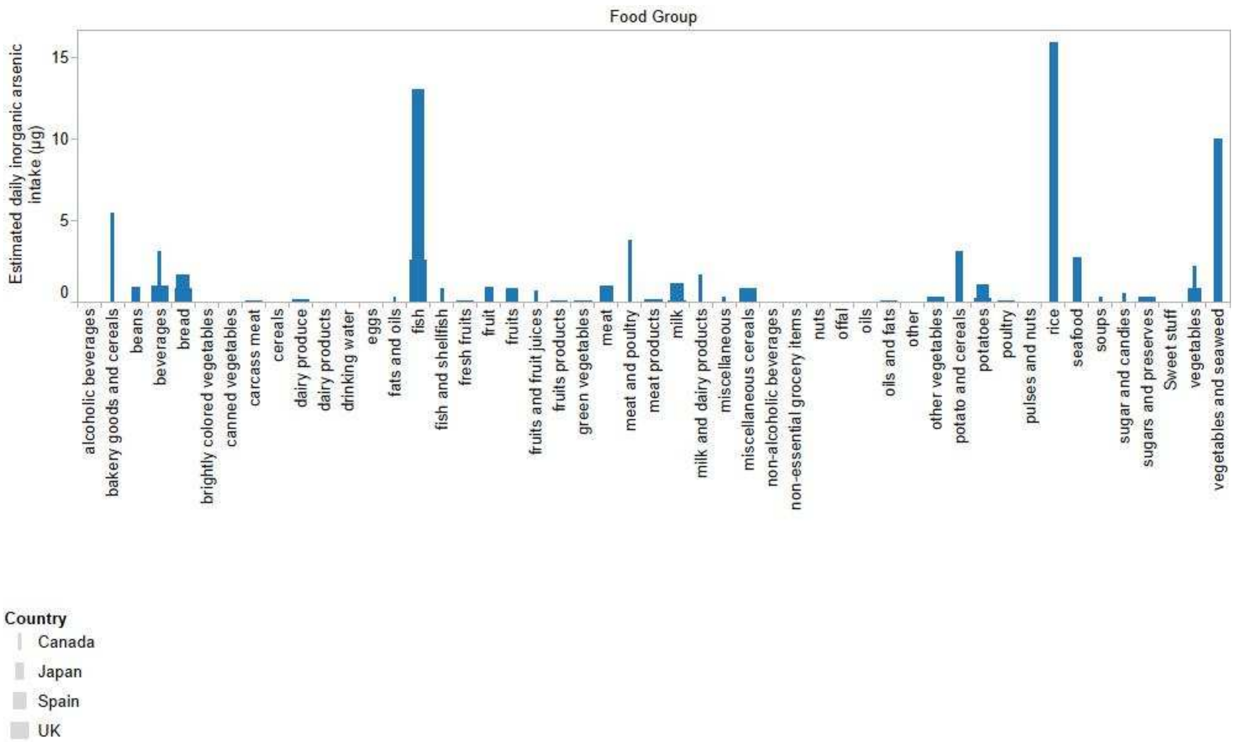


Figure 35. Comparison of estimated daily inorganic arsenic intake for food groups from countries.

Notes: A purpose of this view is to identify food groups reported by multiple countries. In the view, the countries are represented by bars of varying width.
http://public.tableausoftware.com/views/mj_usa_tds/uneyema_tds02

4.3 Objective 3: Compare arsenic content in rice product categories.

4.3.1 Dataset Description

Table 9 presents a summary of the number of entries associated with the data fields in the first analytical results of arsenic content of 193 samples of rice and rice products collected in the U.S. marketplace. There were 5 product categories in the dataset: Basmati Rice, Rice (non-Basmati), Rice Beverage, Rice Cakes and Rice Cereal. A total of 115 Sample Descriptions were reported including Aged, Aged Milled, Basmatic Brown Rice, California Brown, Cracked Wild Rice, Indian Basmati White, Lightly Salted, Long Grain Rice, Medium Grain Rice, Organic Brown, Puffed Rice, Ricemilk Vanilla and Xtra Long Grain Rice. The three countries of origin reported were India, Italy and USA. The country of origin was also designated as "Not Determined" and "Not Reported". The arsenic contents in ppb dry wt were reported as DMA, Inorganic As, MMA and Total Arsenic (Table 9). Additionally, 69 unique values were reported for Inorganic Arsenic per serving (mcg/serving). The values ranged from 1.2 to 11.1.

Table 9. Range and unique values of arsenic species in rice varieties obtained in United States.

Arsenic Species	Range of Content (ppb, dry wt)	Unique Values for Content of Arsenic Species
DMA	12 – 493	103
Inorganic As	14 – 273	112
MMA	0	1
Total Arsenic	5.74 – 723	170

4.3.2 Use Case: Compare rice product categories for Inorganic Arsenic per serving

The dataset was grouped into product categories and then grouped by inorganic arsenic per serving (mcg/serving). A value of 3 mcg/serving was selected as the cut-off since this is equivalent to 10 ppb (limit for arsenic in drinking water set by the U.S. Environmental Protection Agency). The three views (Figures 34 to 36) are respectively for (i) inorganic arsenic up to 2.99 mcg/serving; (ii) inorganic arsenic greater than 3.0 mcg/serving; and (iii) inorganic arsenic not calculated. In the last group the not calculated is speciation not performed or only trace amounts detected (Table 10).

Table 10. Classifying rice samples by a cut-off

Product Category	up to 2.99	>2.99	Not Calculated	Total Samples
Basmati Rice	25	24	3	52
Rice (non-Basmati)	2	47	0	49
Rice Beverage	0	3	25	40
Rice Cake	0	32	0	32
Rice Cereal	12	20	0	32

A total of 39 rice samples were identified in three categories to have inorganic arsenic per serving of up to 2.99 mcg/serving. The categories and number of samples were Basmati Rice [25], Rice (non-Basmati) [2] and Rice Cereal [12]. A total of 126 rice samples were identified in five categories to have inorganic arsenic serving of greater than 2.99 mcg/serving. There categories and number of samples were Basmati Rice [24], Rice (non-Basmati) [47], Rice Beverage [3], Rice Cake [32] and Rice Cereal [20]. A total of 28 rice samples were identified in two categories in which the inorganic arsenic content was not calculated. The categories and number of samples were Basmati Rice [3] and Rice Beverage [25].

Selected samples with lowest and highest values for the two broad categories of inorganic arsenic per serving are presented by Country of Origin, Sample Description and Sample ID (Table 11). For example, the lowest arsenic content observed in Basmati Rice was 1.2 mcg/serving from White Basmati (721854B) from USA.

Table 11. Rice samples categorized by country of origin and inorganic arsenic contact.

	Lowest Content in <2.99 Inorganic Arsenic per serving (mcg/serving)				Highest Content in >2.99 Inorganic Arsenic per serving (mcg/serving)			
Product Category	Country	Arsenic Content	Sample Description	Sample ID	Country	Arsenic Content	Sample Description	Sample ID
Basmati Rice	USA	1.2	White Basmati	721854B	USA	9	Basmati Brown Rice	720110
					USA	9	Brown	721867B
Rice (non-Basmati)	Not Determined	2.7	Long Grain, Enriched Pre-cooked	492965	USA	11.1	Brown Natural Whole Grain	70135
Rice Beverage					Not Reported	4.1	Organic	719976
Rice Cakes					Not Reported	8.2	Salt Free	720115A
					Not Reported	8.2	Sodium Free Plain Gluten Free	725136
Rice Cereal	Not Reported	1.5	Organic Puffs	721859B	Not Reported	9.7	Gluten Free Cream of Brown Rice	719981
	Not Reported	1.5	Whole Grain Puffed Rice, whole grain brown	721862B				

Notes: The lowest arsenic content observed in Basmati Rice was 1.2 mcg/serving from White Basmati (721854B) from USA. In the same product category, a Basmati Brown Rice with Sample ID 720110 from USA had the highest value of 9 mcg/serving.

Product Category	Inorganic As per serving (mcg/serving)	Country of Origin	Sample Description	Sample ID	
Basmati Rice	1.2	USA	White Basmati	721854B	1
	1.8	India	Basmati Rice	720015	1
				720117A	1
	1.9	India	Indian Basmati Aged	721865R	1
	2.1	India	Basmati Rice	720011	1
				720021	1
				720129	1
				728650	1
			Boil-in-Bag	728652	1
	2.2	India	Pure Basmati Rice	720022	1
	2.3	India	Basmati White, Boil-in-bag	492964	1
			Boil-in-Bag	721875	1
		Not Determined	Basmati Rice	70141	1
		USA	California White Organic	720126	1
			Organic White Basmati	720109A	1
				720114	1
			Short Grain White	721863B	1
			White Basmati	720012	1
	2.4	India	Basmati Rice	720111	1
				720128	1
		Not Determined	Basmati Rice	70140	1
	2.5	India	Basmati Rice	720127	1
		USA	Organic White Basmati	720109B	1
	2.9	India	Aged Milled	492966	1
			Indian Basmati White	721864B	1
Rice (non-Basmati)	2.2	Not Determined	Instant Enriched Long Grain	725138	1
	2.7	Not Determined	Long Grain, Enriched Pre-cooked	492965	1
Rice Cereal	1.5	Not Reported	Organic Puffs	721859B	1
			Whole Grain Puffed Rice, whole grain brown rice	721862B	1
	1.6	Not Reported	Whole Grain Puffed Rice, whole grain brown rice	242775	1
	1.7	Not Reported	Rice Cereal	242771	1
	1.8	Not Reported	Puffed Rice	725133	1
	1.9	Not Reported	Puffed Brown Rice	719987	1
			Rice Squares, Crispy Toasted Gluten Free	719988	1
	2	Not Reported	Rice Single Grain (infant)	719983	1
	2.3	Not Reported	Fruity Dyno-Bites	719990	1
	2.6	Not Reported	Cream of Rice	719986	1
	2.9	Not Reported	Crispy Rice Gluten Free Whole Grain Brown	70147	1
			Organic Whole Grain Rice Cereal (infant)	720341	1

Figure 36. Rice samples with inorganic arsenic per serving =<2.9

Notes: http://public.tableausoftware.com/views/mj_usfda_arsenic_rice/conc_lt3

Product Category	Inorganic As per serving (mcg/serving)	Country of Origin	Sample Description	Sample ID	
Basmati Rice	3	USA	Organic Brown Basmati	242778	1
	3.1	India	Aged	728649	1
	3.2	India	Imported Aged	721863A	1
	3.3	India	White Basmati Rice	728645	1
	3.5	India	White Basmati Rice	728653	1
	3.7	India	Basmati Rice	720020	1
			Quick-Cook Organic Brown, parboiled	728646	1
	3.9	India	Organic Aged	721867	1
			Organic White Basmati Rice	721857B	1
			White Basmati Rice	721858B	1
		USA	California Brown	720013	1
	4	India	Organic	728644	1
	4.1	USA	California Brown	720130	1
	4.7	Not Determined	Organic White Basmati	721866	1
		USA	Organic Brown Basmati	720112	1
	4.8	USA	Basmati Rice	720019	1
	5.3	India	Brown Basmati Rice	728654	1
	5.6	India	Brown Basmati Rice	728647	1
	5.7	India	Brown Basmati Rice	721855B	1
	6	India	Brown	720016	1
	6.5	India	Basmati Rice	720018	1
	6.6	USA	Organic Long Grain Brown Basmati	720014	1
	9	USA	Basmati Brown Rice	720110	1
			Brown	721867B	1
Rice (non-Basmati)	3	USA	Instant Enriched Long Grain	492961	1
	3.2	Not Determined	Long Grain White Rice Fully Cooked, parboiled	728638	1
	3.3	Not Determined	Ready to Serve Long Grain White Rice Fully Cooked, parboiled	492963	1
	3.6	Not Determined	Whole Grain Brown Pre-cooked parboiled	725130	1
	4	USA	Whole Grain Red	721856A	1
	4.2	Italy	Organic Carnaroli	721866B	1
	4.4	USA	Medium Grain Rice	720004	1
	4.7	Not Determined	Enriched Rice Long Grain	725139	1
	5.1	USA	Long Grain Brown	721857	1
	5.4	Not Determined	Enriched Long Grain	720000	1
		USA	Extra Long Grain	720003	1
			Long Grain Brown	721860B	1
	5.8	USA	Long Grain White	70139	1
	5.9	USA	Enriched Long Grain Rice	720005	1
			Extra Long Grain Rice	719998	1
			Wild Rice	721856B	1
	6	USA	Wild Riced	721868A	1
	6.2	USA	Cracked Wild Rice	721868B	1
			Long Grain Rice	242774	1
				719997	1
			Medium Grain	720007	1
	6.5	USA	Enriched Extra Long Grain	492959	1
	6.6	USA	Extra Long Grain	70137	1
			Whole Grain Brown Rice	720122	1
	6.8	Not Determined	Long Grain Rice	242777	1
		USA	Extra Long Grain	720008	1
				721877	1
	6.9	USA	Extra Long Grain Rice	721869	1
	7.1	USA	Xtra Long Grain Rice	720124	1
	7.4	USA	Enriched Parboiled Long Grain	492960	1
			Long Grain	70138	1
	7.6	Not Determined	Brown Rice Natural Long Grain	492962	1
	7.9	USA	Whole Grain Brown Rice	720123	1
	8	USA	Extra Long Grain	720010	1
			Whole Grain Brown	70134	1
	8.2	USA	Brown Rice Whole Grain	720125	1
	8.4	USA	Brown Natural Whole Grain	719999	1
	8.8	USA	Whole Grain, Long Grain	492958	1
	9.1	USA	Long Grain Brown Whole	720009	1
	9.4	USA	Long Grain Brown Rice Extra Fancy	721864A	1
			Whole Grain Brown Rice	720120	1
	9.7	Not Determined	Brown Rice Natural Long Grain	725137	1
	10.2	Not Determined	Long Grain Brown Rice	720006	1
	10.4	USA	Long Grain Brown Rice	70136	1
			Whole Grain Brown Rice	720121	1
	10.5	USA	Long Grain Natural Brown Rice	720001	1
	11.1	USA	Brown Natural Whole Grain	70135	1
Rice Beverage	3.4	Not Reported	Rice Drink Organic Original	721854A	1
	3.8	Not Reported	Organic	719980	1
	4.1	Not Reported	Organic	719976	1
Rice Cakes	3	Not Reported	Organic Unsalted	719993	1
				719994	1
			Wild Organic Gluten-free Vegan Lightly Salted	721871	1
	3.1	Not Reported	Brown Lightly Salted	728639	1
			Caramel Corn Gluten-free	720119	1
			Organic Brown	721878	1
	3.2	Not Reported	Organic Brown	728640	1
	3.7	Not Reported	Brown Rice Organic Salt-free	719991	1
	3.8	Not Reported	Brown Rice Organic Salt-free	721858	1
	4	Not Reported	Butter Popped Corn	720116	1
	4.6	Not Reported	Lightly Salted	721876	1
	4.7	Not Reported	Brown Rice Whole Grain Lightly Salted	242776	1
	4.8	Not Reported	Brown Salt Free	721862A	1
	5	Not Reported	Apple Cinnamon	720118	1
	5.1	Not Reported	Apple Cinnamon	719997B	1
	5.6	Not Reported	Butter Popped Corn	720117B	1
			Salt Free	721865B	1
	6.1	Not Reported	Salt Free	728642	1
	6.2	Not Reported	Lightly Salted	719996	1
			Salt Free	70133	1
	6.3	Not Reported	Salt Free	70131	1
				70132	1
	6.5	Not Reported	Plain Salt Free Whole Grain Brown	728641	1
	6.6	Not Reported	Plain Salt Free Whole Grain Brown	725135	1
	6.8	Not Reported	Salt Free, whole grain brown rice	721865	1
	7.2	Not Reported	Low Sodium	719995	1
	7.6	Not Reported	Salted Plain Gluten Free	728643	1
	7.7	Not Reported	Lightly Salted	719992	1
	8	Not Reported	Salt Free	242773	1
				720115B	1
	8.2	Not Reported	Salt Free	720115A	1
			Sodium Free Plain Gluten Free	725136	1
Rice Cereal	3.2	Not Reported	Crispy Brown Rice Organic	492957	1
			Crispy Rice Gluten Free Whole Grain Brown	70146	1
			Crispy Rice Gluten-free Whole Grain Brown Rice	725131	1
			Crispy Rice Toasted	70142	1
			Organic Whole Grain Rice Cereal (infant)	70145	1
	3.4	Not Reported	Crispy Brown Rice Gluten Free Organic	721859	1
	3.6	Not Reported	Crispy Rice Toasted	725134	1
	3.8	Not Reported	Crispy Rice	719989	1
			Rice Cereal	242770	1
	3.9	Not Reported	Crispy Rice	719984	1
			Crispy Rice Toasted	70143	1
				70149	1
			Crispy Toasted Rice	719985	1
	4.1	Not Reported	Crisp Rice Toasted	721869B	1
			Honey Rice Gluten Free	719982	1
	4.2	Not Reported	Crispy Rice Toasted	721874	1
	4.3	Not Reported	Rice Chex Gluten-free	725132	1
	7.1	Not Reported	Organic Brown Rice Crisps	70144	1
	7.3	Not Reported	Organic Brown Rice Crisps	720342	1
	9.7	Not Reported	Gluten Free Cream of Brown Rice	719981	1

Figure 37. Rice samples with inorganic arsenic per serving >2.9

http://public.tableausoftware.com/views/mj_usfda_arsenic_rice/conc_gt3

Product Category	Inorganic As per serving (mcg/serving)	Sample Description	Sample ID	
Basmati Rice	Not Calculated	Fully Cooked	728648	1
		Pure Basmati Rice	720113	1
		Whole Grain Organic Microwave	728651	1
Rice Beverage	Not Calculated	Enriched Rice Beverage	720333	1
			720339	1
		Lower Fat Plain Rice Beverage	720340	1
		Organic	242772	1
			719977	1
			719978	1
			719979	1
			720334	1
			720335	1
		Organic Rice Beverage	720338	1
		Organic Unsweetened Rice Beverage	721861 B	1
		Organic Unsweetened Rice Beverage2	720337	1
		Organic Vanilla Rice Beverage	70148	1
		Organic Whole Grain Rice Beverage	720336	1
		Original Rice Beverage	70154	1
			70155	1
		Rice Beverage	721870	1
		Rice Beverage Original	70151	1
			70152	1
			70156	1
		Rice Beverage Plain	721860A	1
		Rice Drink Original Classic	721861	1
		Rice Drink Original Enriched	721855	1
		Ricemilk Vanilla	70153	1
		Vanilla Rice Beverage	70150	1

Figure 38. Rice samples with negligible inorganic arsenic per serving.

http://public.tableausoftware.com/views/mj_usfda_arsenic_rice/NotCalculated

4.4 Objective 4: Identify informative sentences on arsenic concentrations in rice.

4.4.1 Dataset and View Description

A total of 86 sentences mapped to 53 PubMed abstracts were identified to be informative for arsenic concentrations from the 758 sentences (240 PubMed abstracts) obtained from the Arsenic Sentence Database [11]. The complete list of 86 sentences is

[11] http://genomics.jsums.edu/sentence/arsenic_pubmed/

presented in Appendix 3. In the view, the field were arranged as follows: PMID, Sentence Identifier and Sentence Text. Filters for the fields were also included in the view. A subset of 31 sentences from 21 PubMed abstracts contained the word 'concentration' (Figure 39). An insight from the set of sentences was on limits of arsenic in rice. One of the PubMed abstracts PMID 18546734 (Zavala Duxbury 2008) reported a global "normal" range of 0.08-0.20 mg kg(-1). This estimate was obtained from comparing 204 commercial retail rice from mostly upstate New York and supplemented with samples from other countries.

The concentration of arsenic, sample source, arsenic species, year of publication was manually extracted from the sentences with the aid of a dashboard (Figure 40) displayed the sentence text and a webpage for the PubMed abstract corresponding to the sentence. The extracted information was compiled in a spreadsheet file and then uploaded into the visual analytics software. A visual analytics task completed was to integrate the arsenic concentration in rice investigated the selected PubMed abstracts. The complete arsenic concentration table is available in Appendix 4. A subset of the table for PubMed abstracts that included concentrations of arsenic in rice from Bangladesh is presented in Table 12. An integration of the data is presented in Figure 41.

PMID	Sentence Identifier	Sentence Text	
11918027	11918027_8	Concentrations of arsenic in rice grain did not exceed the food hygiene concentration limit (1.0 mg of As kg(-1) dry weight).	1
	11918027_9	The concentrations of arsenic in rice straw (up to 91.8 mg kg(-1) for the highest As treatment) were of the same order of magnitude as root arsenic concentrations (up to 107.5 mg kg(-1)), suggesting that arsenic can be readily translocated to the shoot.	1
12059152	12059152_6	Their levels were, in general, below the maximum limits establish by the Spaniard legislation; however, the As concentration in the licorice sticks was above this maximum limit (0.11 +/- 0.01 microg g(-)(1)).	1
	12059152_9	The As concentration in the licorice extract was 0.503 +/- 0.01 microg g(-)(1), and could represent a serious hazard to human health if it is used in high proportions.	1
12564892	12564892_10	Rice grain grown in the regions where arsenic is building up in the soil had high arsenic concentrations, with three rice grain samples having levels above 1.7 microg g(-1).	1
14987870	14987870_7	However, rice plants, especially the roots had a significantly higher concentration of arsenic (2.4 mg/kg) compared to stem (0.73 mg/kg) and rice grains (0.14 mg/kg).	1
15084107	15084107_5	Arsenic concentrations found in rice-based cereals (63-320 ng/g dry weight) were similar to those reported for raw rice.	1
15979720	15979720_6	High As concentrations were found in the rice grain that ranged from 0.5 to 7.5 mg/kg, most of which exceed the maximal permissible limit of 1.0 mg/kg dry matter.	1
16003581	16003581_4	The concentration of the total arsenic in the samples i.e. water (n = 64), soil (n = 30), sediment (n = 27) and rice grain (n = 10) were ranged from 15 to 825 microg L(-1), 9 to 390 mg kg(-1), 19 to 489 mg kg(-1) and 0.018 to 0.446 mg kg(-1), respectively	1
16082952	16082952_9	The dimethylarsinic acid (DMA) and inorganic arsenic concentration ranged from 22 to 270 ng of As/g of rice and from 31 to 108 ng of As/g of rice, respectively, for samples cooked in reagent water.	1
16730050	16730050_6	Digestion and analysis of individual grains of boro winter rice from the 2 sites irrigated with groundwater containing 150 and 180 microg/L As yielded concentrations of 0.28+/-0.13 mg/kg (n=12) and 0.44+/-0.25 mg/kg (n=12), respectively.	1
16839594	16839594_7	Arsenic concentrations were 0.40+/-0.03 and 0.58+/-0.12 mg/kg in parboiled rice of arsenic affected area, cooked with excess water and 1.35+/-0.04 and 1.59+/-0.07 mg/kg in gruel for BRRI dhan28 and BRRI hybrid dhan1, respectively.	1
	16839594_8	In non-parboiled rice, arsenic concentrations were 0.39+/-0.04 and 0.44+/-0.03 mg/kg in rice cooked with excess water and 1.62+/-0.07 and 1.74+/-0.05 mg/kg in gruel for BRRI dhan28 and BRRI hybrid dhan1, respectively.	1
17067657	17067657_3	In the unaffected areas, where irrigation water contained little As (<1 microg/L), As concentrations of rice field soils ranged from 1.5 to 3.0 mg/kg and did not vary significantly with either depth or sampling time throughout the irrigation period.	1
17346792	17346792_6	Arsenic concentration in rice grain was 0.5+/-0.02 mg kg(-1) with the highest concentrations being in grains grown on soil treated with 40 mg As kg(-1) soil.	1
	17346792_9	Moreover, when the rice plant was grown in 60 mg of As kg(-1) soil, arsenic concentrations in rice straw were 20.6+/-0.52 at panicle initiation stage and 23.7+/-0.44 at maturity stage, whereas it was 1.6+/-0.20 mg kg(-1) in husk.	1
17599387	17599387_8	Arsenic concentrations in parboiled and non-parboiled brown rice of BRRI dhan28 were 0.8+/-0.1 and 0.5+/-0.0 mg kg(-1) dry weight, respectively while those of BRRI hybrid dhan1 were 0.8+/-0.2 and 0.6+/-0.2 mg kg(-1) dry weight, respectively.	1
	17599387_11	The concentration of arsenic found in the present study is much lower than the permissible limit in rice (1.0 mg kg(-1)) according to WHO recommendation.	1
17852383	17852383_6	The results show that the arsenic concentration in cooked rice is always higher than that in raw rice and range from 227 to 1642 microg kg(-1).	1
	17852383_9	The daily inorganic arsenic intakes for water plus rice were 229, 1024 and 2000 microg day(-1) for initial arsenic concentrations in the cooking water of 50, 250 and 500 microg arsenic l(-1), respectively, compared with the tolerable daily intake which is	1
18385862	18385862_3	All rice milk samples analysed in a supermarket survey (n = 19) would fail the EU limit with up to 3 times this concentration recorded, while out of the subset that had arsenic species determined (n = 15), 80% had inorganic arsenic levels above 10 microg	1
18503245	18503245_9	When the method was applied to ten short-grain brown rice samples, the iAs concentrations were 0.108-0.227 mg/kg dry weight and the total As concentrations were 0.118-0.260 mg/kg dry weight.	1
18546734	18546734_5	Total As concentration in rice varied from 0.005 to 0.710 mg kg(-1).	1
	18546734_6	We combined our data set with literature values to derive a global "normal" range of 0.08-0.20 mg kg(-1) for As concentration in rice.	1
	18546734_7	The mean As concentrations for rice from the U.S. and Europe (both 0.198 mg kg(-1)) were statistically similar and significantly higher than rice from Asia (0.07 mg kg(-1)).	1
	18546734_9	Wide variability found in U.S. rice grain was primarily influenced by region of growth rather than commercial type, with rice grown in Texas and Arkansas having significantly higher mean As concentrations than that from California (0.258 and 0.190 versus	1
18546736	18546736_6	After cooking, the rice retained between 45% and 107% of the As(V) added to the cooking water, and the inorganic As concentrations ranged between 0.428 microg g(-1) dw (0.1 microg mL(-1) in the cooking water) and 3.89 microg g(-1) dw (1.0 microg microL(-1	1
18637329	18637329_11	Arsenic concentrations in rice grain are lower than the food safety limitation in China (0.7 mg x kg(-1)).	1
18678041	18678041_8	The rice collected from mine-impacted regions, however, were found to be highly enriched in As, reaching concentrations of up to 624 ng g(-1).	1
	18678041_11	The mean baseline concentrations for As(i) in Chinese market rice based on this survey were estimated to be 96 ng g(-1) while levels in mine-impacted areas were higher with ca. 50% of the rice in one region predicted to fail the national standard.	1
19948354	19948354_3	Two-week-old rice seedlings were exposed to two concentrations of arsenate (50 or 100 microM), and leaf samples were collected 4d after treatment.	1

Figure 39. Informative sentences on arsenic concentration in rice.

Notes: http://public.tableausoftware.com/views/mj_rice_arsenic_sentences/rice_arsenic_sentences

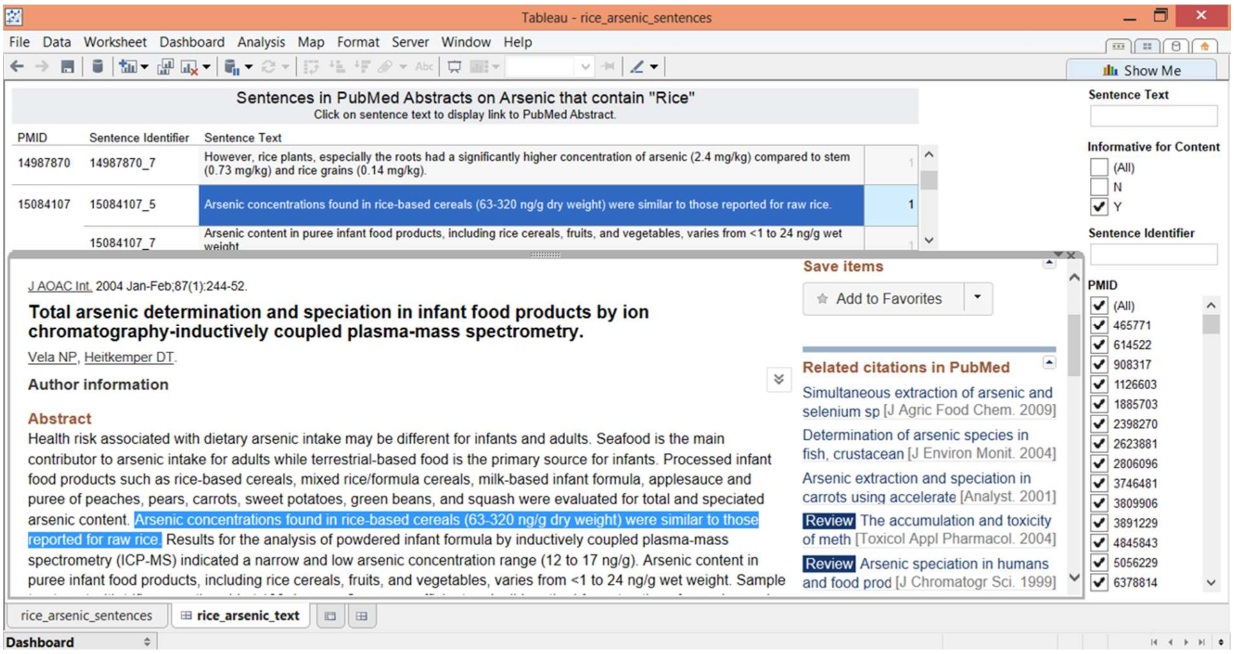


Figure 40. Dashboard for connecting sentences to literature knowledgebase.

Notes: http://public.tableausoftware.com/views/mj_rice_arsenic_sentences/rice_arsenic_text#1

Location	PMID	Year of Publication	Sample	Arsenic Species	Arsenic Content
Bangladesh	11918027	2002	Rice straw	Arsenic	up to 91.8 mg kg(-1)
	12564892	2003	Rice grains	Arsenic	>1.7 micro g(-1)
	14987870	2004	Rice grains	Arsenic	0.14 mg/kg
			Rice roots	Arsenic	2.4 mg/kg
			Rice stem	Arsenic	0.73 mg/kg
	16730050	2006	Boro Winter Rice irrigated with groundwater containing 150 microg/L arsenic	Arsenic	0.28+/-0.13 mg/kg
			Boro Winter Rice irrigated with groundwater containing 180 microg/L arsenic	Arsenic	0.44+/-0.25 mg/kg
	16839594	2006	Non-Parboiled Rice Gruel for BRRI dhan28	Arsenic	1.62+/-0.07 mg/kg
			Non-Parboiled Rice Gruel for BRRI hybrid dhan1	Arsenic	1.74+/-0.05 mg/kg
			Non-Parboiled Rice of Affected Area for BRRI dhan28, cooked with excess water	Arsenic	0.39+/-0.04 mg/kg
			Non-Parboiled Rice of Affected Area for BRRI hybrid dhan1, cooked with excess water	Arsenic	0.44+/-0.03 mg/kg
			Parboiled Rice Gruel for BRRI dhan28	Arsenic	1.35+/-0.04 mg/kg
			Parboiled Rice Gruel for BRRI hybrid dhan1	Arsenic	1.59+/-0.07 mg/kg
			Parboiled Rice of Affected Area for BRRI dhan28, cooked with excess water	Arsenic	0.40+/-0.03 mg/kg
			Parboiled Rice of Affected Area for BRRI hybrid dhan1, cooked with excess water	Arsenic	0.58+/-0.12 mg/kg
	17346792	2008	Hust of rice plant grown on soil treated with 60 mg As kg(-1)	Arsenic	1.6+/-0.20 mg kg(-1)
			Rice grains grown on soil treated with 40 mg As kg(-1)	Arsenic	0.5+/-0.02 mg kg(-1)
			Straw of rice plant at maturity stage grown on soil treated with 60 mg As kg(-1)	Arsenic	23.7+/-0.44 mg kg(-1)
			Straw of rice plant at panicle initiation stage grown on soil treated with 60 mg As kg(-1)	Arsenic	20.6+/-0.52 mg kg(-1)
	17599387	2007	Non-Parboiled brown rice of BRRI dhan1	Arsenic	0.6+/-0.2 mg kg(-1) dry weight
			Non-Parboiled brown rice of BRRI dhan28	Arsenic	0.5+/-0.0 mg kg(-1) dry weight
			Parboiled brown rice of BRRI dhan1	Arsenic	0.8+/-0.2 mg kg(-1) dry weight
			Parboiled brown rice of BRRI dhan28	Arsenic	0.8+/-0.1 mg kg(-1) dry weight

Figure 41. Integration of data extraction from informative sentences on arsenic content in rice.

http://public.tableausoftware.com/views/mj_rice_arsenic_sentences/rice_conc_pubmed#1

An insight from the grid view is the identification of two research investigations (PubMed Identifiers 16839594 and 17599387) on the arsenic accumulation of dhan1 and dhan28 cultivated rice varieties (Rahman et al., 2006, Rahman et al., 2007). Another investigation (PMID: 14987870) showed the differences in the arsenic concentration of grains (0.14 mg/kg), roots (2.4 mg/kg) and stem (0.73 mg/kg) (Das et al., 2004).

Table 12. Comparison of arsenic content from rice samples from Bangladesh

PMID	Sample	Location	Arsenic Content	Year of Publication	Arsenic Species
11918027	Rice straw	Bangladesh	up to 91.8 mg kg(-1)	2002	Arsenic
12564892	Rice grains	Bangladesh	>1.7 micro g(-1)	2003	Arsenic
14987870	Rice roots	Bangladesh	2.4 mg/kg	2004	Arsenic
14987870	Rice stem	Bangladesh	0.73 mg/kg	2004	Arsenic
14987870	Rice grains	Bangladesh	0.14 mg/kg	2004	Arsenic
16730050	Boro Winter Rice irrigated with groundwater containing 150 microg/L arsenic	Bangladesh	0.28+/-0.13 mg/kg	2006	Arsenic
16730050	Boro Winter Rice irrigated with groundwater containing 180 microg/L arsenic	Bangladesh	0.44+/-0.25 mg/kg	2006	Arsenic
16839594	Parboiled Rice of Affected Area for BRRI dhan28, cooked with excess water	Bangladesh	0.40+/-0.03 mg/kg	2006	Arsenic
16839594	Parboiled Rice of Affected Area for BRRI hybrid dhan1, cooked with excess water	Bangladesh	0.58+/-0.12 mg/kg	2006	Arsenic
16839594	Parboiled Rice Gruel for BRRI dhan28	Bangladesh	1.35+/-0.04 mg/kg	2006	Arsenic
16839594	Parboiled Rice Gruel for BRRI hybrid dhan1	Bangladesh	1.59+/-0.07 mg/kg	2006	Arsenic
16839594	Non-Parboiled Rice of Affected Area for BRRI dhan28, cooked with excess water	Bangladesh	0.39+/-0.04 mg/kg	2006	Arsenic
16839594	Non-Parboiled Rice of Affected Area for BRRI hybrid dhan1, cooked with excess water	Bangladesh	0.44+/-0.03 mg/kg	2006	Arsenic
16839594	Non-Parboiled Rice Gruel for BRRI dhan28	Bangladesh	1.62+/-0.07 mg/kg	2006	Arsenic
16839594	Non-Parboiled Rice Gruel for BRRI hybrid dhan1	Bangladesh	1.74+/-0.05 mg/kg	2006	Arsenic
17599387	Parboiled brown rice of BRRI dhan28	Bangladesh	0.8+/-0.1 mg kg(-1) dry weight	2007	Arsenic
17599387	Non-Parboiled brown rice of BRRI dhan28	Bangladesh	0.5+/-0.0 mg kg(-1) dry weight	2007	Arsenic
17599387	Parboiled brown rice of BRRI dhan1	Bangladesh	0.8+/-0.2 mg kg(-1) dry weight	2007	Arsenic
17599387	Non-Parboiled brown rice of BRRI dhan1	Bangladesh	0.6+/-0.2 mg kg(-1) dry weight	2007	Arsenic
17346792	Rice grains grown on soil treated with 40 mg As kg(-1)	Bangladesh	0.5+/-0.02 mg kg(-1)	2008	Arsenic
17346792	Straw of rice plant at panicle initiation stage grown on soil treated with 60 mg As kg(-1)	Bangladesh	20.6+/-0.52 mg kg(-1)	2008	Arsenic
17346792	Straw of rice plant at maturity stage grown on soil treated with 60 mg As kg(-1)	Bangladesh	23.7+/-0.44 mg kg(-1)	2008	Arsenic
17346792	Hust of rice plant grown on soil treated with 60 mg As kg(-1)	Bangladesh	1.6+/-0.20 mg kg(-1)	2008	Arsenic

CHAPTER 5

DISCUSSION

Arsenic is an environmental toxicant associated with cancer and other chronic diseases including diabetes (Moon et al., 2013, Pan et al., 2013) and possibly chronic kidney disease (Jayasumana et al., 2013, Jayatilake et al., 2013). In humans, drinking water and dietary intake constitute the major route of exposure to arsenic (Halder et al., 2013, Melkonian et al., 2013). The monitoring of arsenic levels in various foods is a priority of health organizations and food safety organizations at global, continental and national levels (Moy, 2013). Furthermore, research investigation on the concentrations of arsenic in various foods is an active area of research (Jackson et al., 2012). Taken together, these monitoring and research activities of arsenic in foods produce volumes and variety of datasets needed to inform the risks associated with consuming certain foods (Chen and Gao, 1993, D'Amato et al., 2013, Egan, 2013, Egan et al., 2007, Egan et al., 2002, FSAI, 2011, Gunderson, 1995, Tao and Bolger, 1999, Wong et al., 2013).

This PhD research has utilized both research datasets and data collected by food safety organizations to make the case for visual analytics of arsenic concentrations in food. This doctoral research in Environmental Science has addressed the need to provide visual analytical decision support tools on arsenic content in various foods. A special emphasis of analysis was on rice, a staple crop in many countries that has received significant attention regarding arsenic content (Gilbert-Diamond et al., 2011, Melkonian et al., 2013, Rahman et al., 2007, Wei et al., 2013, Williams et al., 2007b). The hypothesis of this PhD research is that software enabled visual representation and user interaction facilitated by visual interfaces will help discover hidden relationships between arsenic content and food categories. The visual analytics software enabled the researcher to interact with the datasets to arrive at views that are discussed.

The following four objectives were used to test the hypothesis: (i) provide insightful visual analytic views of compiled data on arsenic in food categories; (ii) categorize table ready foods by arsenic content; (ii) compare arsenic content in rice product categories; and (iv) identify informative sentences on arsenic concentrations in rice. The key results of the research objectives is discussed in this Chapter in the context of the visual analytics tasks assisting the author to discover hidden relationships between arsenic content and food groups.

The focus of Objective 1 of this research was to provide insightful visual analytic views of compiled data on arsenic in foods. The compilation of arsenic concentrations in various crops by Uneyema et al (2007) provided the data source for visual analytics tasks utilized to accomplished Objective 1. Three views (Figure 27 to Figure 29) were constructed to provide insights on arsenic contents of crops and in particular rice.

In Figure 27 with a use case on inorganic arsenic assessment in data, the view reveal two reports from the USA conducted in 2000 and 2004 (Ackerman et al., 2005, Heitkemper et al., 2001) that used ICP-MS to determine the arsenic content in long grain white, long grain brown, wild rice, rice flour and cooked rice. The inorganic arsenic content ranged from 0.0231 to 0.108 ppm. Among the rice varieties investigated, the R5 instant long grain white rice had the lowest arsenic concentration of 0.0231 compared with the other varieties that had ~0.1 ppm. This relatively low arsenic content of instant rice was also observed in the results of close to 200 rice samples analysed in 2012 by the U.S. FDA (Objective 3; Figure 36).

The inorganic arsenic content in the two samples; 492961 and 725138 were 0.066 ppm and 0.049 ppm respectively. Instant (quick) rice is pre-cooked and dehydration in the factory (Luh, 1991, Smith et al., 1985). This pre-processing could account for the low arsenic content compared with other rice products that are not pre-cooked and dehydrated. In summary, the view in Figure 27 allowed for the identification of research articles where the concentrations of inorganic arsenic were determined for crops. In particular two studies that used same method of arsenic speciation were identified. Finally, the outlier arsenic concentration of the instant rice led to suggested basis for the low arsenic content.

Additional insight was on the concentration of dimethylarsinic acid (DMA) rice grains. This distribution of DMA in the rice grain was confirmed in Figure 26. However, an outlier concentration of 0.0099 ppm was observed for Dimethylarsinic Acid (DMA) in the R4 wild rice. In rice grain, DMA(V) is the main form of methylated As, and can account for up to 80% of the total As (Meharg, 2004, Zavala and Duxbury, 2008, Zavala et al., 2008). Wild rice is actually an aquatic grass of the genus *Zizania* (Dong et al., 2013). This lower accumulation of DMA in rice grain could be due to differences in translocation of arsenic by *Oryza* species (true rice) and aquatic grasses. The translocation of dimethylarsinic acid to the rice grain was over an order of magnitude higher than for inorganic species (Carey et al., 2010).

Another use case grouped rice samples by country of origin. In constructing this view, the visual analytics tasks included data transformation by identifying research investigations on total arsenic and subsequently extracting only the maximum concentration reported (Figure 28). This view both provided insights on (i) cluster of research according to area (country) where the research was done; and (ii) differences in arsenic concentrations reported. Outlier concentrations of 2.590 ppm was reported for Malaysia (Zarcinas et al., 2004). This outlier data provided interest to identify recent research results on arsenic concentrations of rice grown in Malaysia (Moon et al., 2012, Salim et al., 2010). Salim et al. (2010) determined by instrumental neutron activation analysis the elemental arsenic concentrations (mg/kg or ppm)) in 12 types of rice: Village Rice (0.11+/-0.02); Local Rice (0.11+/-0.01); Thailand Rice (0.15+/-0.01); Basmathi Rice (0.08+/-0.01); Hill Rice (0.10+/-0.02); Bario Keladi Rice (0.13+/-0.01); Red Bario Rice A (0.16+/-0.01); White Bario Rice (0.10+/-0.01); Red Bario Rice B (0.12+/-0.01); Black Glutinous Rice (0.13+/-0.01); Husk Open Rice (0.19+/-0.03) and Brown Rice (0.19+/-0.01). The mean, minimum and maximum concentrations of inorganic arsenic for the 12 rice types obtained in Malaysia were 0.13, 0.08 and 0.19 mg/kg respectively. In another report, comparing white rice in 7 Asian countries, the elemental arsenic content obtained for white rice from Malaysia was 0.11 mg/kg (Moon et al. 2010). Another interesting data from Moon et al. (2010) was the daily intake values (µg) for arsenic via white rice: China (165 µg); Indonesia (24 µg); Japan (31 µg); Korea (39 µg); Malaysia (33 µg); Philippines (21 µg) and Thailand (27 µg).

The focus of Objective 2 of this research was to categorize table ready foods in total diet studies into food groups and subsequently visual analytics techniques were used to explore data to determine relationships between arsenic and food groups. In this objective, three use cases were assessed on the USA Total Diet Study on 328 table ready foods: (i) identify foods with arsenic; (ii) identify foods without arsenic; (iii) group foods by category and arsenic content. An additional use case was to compare the estimate daily inorganic arsenic intake (µg) for food group from four countries. A 55-chapter book on diverse aspects of total diet studies underscores the importance of total diet studies as the most cost-effective means of assessing the safety and nutritional quality of diets (Moy and Vannoort, 2013). Furthermore, the need has been identified for centralized resources to manage data on chemicals in food and diet from a global perspective (Sommerfeld and Moy, 2013).

The 328 table ready foods were categorized into 12 food categories (baby foods; beverage; dairy; eggs; fats/oils; fruits; grains; legumes; meat, poultry, fish; mixtures; sweet and vegetables. This categorization of the table ready foods is a contribution of this thesis and enhances the communication of the data available on the United States Food and Drug Administration (US FDA) website. This food categorization addresses the need for appropriate strategies to communicating TDS data (Flynn, 2013). The uses cases provide views that communicate arsenic content and dietary exposures in the TDS conducted in the U.S. (Figure 30).

The table ready foods were also categorized into 185 foods with arsenic and 143 foods without arsenic. The combination of categories allowed for identifying table ready foods made from rice that contain arsenic (Figure 31). No rice-containing table ready food was observed in the “foods without arsenic” category. This observation is supported by research that has shown that rice is more efficient than other cereals at accumulating arsenic in grain and shoot (Carey et al., 2010, Jia et al., 2013). These two broad categories (with/without arsenic) could help in decision making by consumers on selection of table ready foods. The view (Figure 33) that group food by category and arsenic content integrates the maximum and mean concentration to determine the arsenic contents in selected groups that include grains. Interestingly, the grouping by maximum concentration also reveals a table ready food in the grain food category that includes rice with a mean arsenic concentration of 0.046 mg/kg.

The visual analytics software enabled integration of TDS data from four countries (Figure 34). The view produced also identified outlier daily inorganic arsenic intake of 15.87 µg for rice. The recommendations for standards in collecting TDS data will enable diverse forms of analysis of data from multiple countries (Mooney et al., 2013, Sommerfeld and Moy, 2013). The view in Figure 35 allows for comparison of estimated daily inorganic arsenic intake values of food groups from multiple countries. In Figure 34 and Figure 35, bar graphs were utilized to encode the data since bar graph are bar graphs are effective for simple judgments of proportion (Hegarty, 2011). Total Diet Studies are currently limited from sub-Saharan African countries (Gimou et al., 2013, Gimou et al., 2008). The use of visual analytics could facilitate collaborative analysis and sense-making on TDS from African countries.

The focus of Objective 3 of the research was to compare arsenic content in rice product categories in 193 rice samples. The September 2012 full analytical results of arsenic in rice by the U.S. FDA provided the data source for objective. In a decision making process on arsenic in foods, developing acceptable levels for arsenic concentration in foods could assist stakeholders and consumers in risk assessment (Consumer Reports, 2012, Consumer Reports, 2013, Rintala et al., 2014). In July 2013, the U.S. FDA set an action level for inorganic arsenic in single-strength apple juice of 10 µg/kg or 10 ppb (Consumer Report 2013). In China, the legal maximum level on inorganic arsenic in rice is 0.15 mg/kg (Sloth). The CODEX Alimentarius Commission has proposed a maximum level of 0.3 mg/kg for raw rice (CODEX, 2012). The statement from CODEX is “draft MLs for As in raw rice (brown) would be proposed at 0.3 mg/kg, whether for inorganic As or total As; or 0.2 mg/kg only for inorganic As in polished rice. It might be measured for total As first, and then measured as inorganic As if the total As measurement exceeds 0.3 mg/kg.”

As the methods for determining arsenic in foods become widely used, there is possibility that datasets on analytical results for several foods including rice and seafood will become available to consumers (CODEX, 2012, Consumer Reports, 2012, Consumer Reports, 2013, JEFCA, 2011, JEFCA, 1995, Kawalek, Lasky, 2013, Nachman et al., 2013b). Considering the amount of data that could become available, stakeholders and consumers would need to explore, mine and analyse data for various purposes and guided by the acceptable arsenic levels.

Thus in this research, a level of 3 microgram/serving for inorganic arsenic was utilized to categorize the 193 rice samples (Table 10 and Table 11). A sample labelled “Instant Enriched Long Grain” (725138) had a value of 2.2 mcg/serving in the dataset with values that ranged from 1.2 to 11.1. This is consistent with observations in Objective 1, where an instant rice species had lower arsenic contents compared with other rice types.

The focus of Objective 4 of this research was to identify informative sentences on arsenic concentrations in rice using a visual analytics enhanced literature curation approach. A total of 86 sentences mapped to 53 PubMed abstracts were identified to be informative for arsenic concentrations from the 758 sentences (240 PubMed abstracts) obtained from the Arsenic Sentence Database. The information in the sentences relating to the arsenic concentration in rice and rice-based products were varied. The information types included: (i) range of arsenic concentrations in cooked and raw rice; (ii) differences in arsenic concentration of based on cooking methods such as limited and excess water; (iii) arsenic concentration in rice from arsenic-affected areas or market basket surveys; and (iv) acceptable concentration of arsenic in rice.

An informative sentence from PubMed abstract 18546734 provided the global “normal” range of 0.08 to 0.20 mg/kg for arsenic in rice. Interestingly, the Codex Alimentarius proposed a maximum inorganic level in raw rice of 0.20 mg/kg (CODEX, 2012, Sloth). The search for informative sentences used the word “concentration” as the filter. This would miss other informative sentences that included the use of levels or content. Further research could integrate multiple search terms.

A focus area of visual analytics is the interaction with data (Sedig and Parsons, 2013, Thomas and Cook, 2005). In Objective 4 of the research, a dashboard was developed to facilitate the interaction with text and a connection to the knowledge base of the PubMed literature database (Figure 40). The dashboard facilitated the extraction of facts from the informative sentences which were then compiled into a data table (Table 12; Appendix 5). The data table was explored using the study location information obtained from PubMed abstract. The example of Bangladesh provided insights on several concentrations reported in the research articles. Interactive dashboard designs have benefited from collaboration between a visual analytics expert and a subject matter expert (Al-Hajj et al., 2013). This research provides basis for future integration of semantic information with quantitative data. Further research on the development of dashboards for arsenic in foods could include extensive user evaluation which is outside the scope of this thesis.

In summary, the four objectives led to visual analytics facilitated discoveries of hidden relationships between arsenic content and food categories. As massive amounts and variety of data are envisioned for arsenic in foods, complex cognitive tasks would need to be accomplished. These tasks include problem solving, decision making, sense making, learning and analysing. Visual analytics facilitated exploration, mining and analysis would help regulatory agencies, manufacturers, consumers and other stakeholders to accomplish these complex cognitive tasks concerning arsenic contents of foods. Finally, care must be taken with interpretation of the results of data visualization patterns as data used may be incomplete (Vincent et al., 2010).

CHAPTER 6

CONCLUSIONS AND RECOMENDATIONS

6.1 Conclusions

Several datasets were utilized to make the case for visual analytics of data on arsenic concentrations in foods. A secondary data analysis was the overall approach for the research. The dataset on crops provided by Uneyema et al (2007) consisted of 459 crop samples with 9 dimensions (Area, Content, Method, Molecular Species, ppm (parts per million), References, Sample Number, Species and Year). The second dataset consisted of consisted of 328 table ready foods and 10 data fields for arsenic content in the Total Diet Study (March 1991 to April 2005). The data fields (table columns) were TDS Food Description, TDS Food No., Number of Results, Number Not Detected, Number of Traces, Mean (mg/kg), Std Dev [Standard Deviation] (mg/kg), Minimum (mg/kg), Maximum (mg/kg) and Median (mg/kg).

An integrated dataset of estimates of daily inorganic arsenic intake was constructed for 49 food groups from TDS from four countries (Canada, Japan, Spain and United Kingdom). A total of 31 unique estimates were observed in the dataset. A dataset utilized to compare arsenic content in rice product categories for 193 samples of rice and rice products that were further divided into 5 product categories and mapped to sample descriptions; sample identifiers, country of origin, content of arsenic species, and inorganic arsenic per serving. Finally, the dataset from PubMed literature consisted of 758 sentences on arsenic in rice from 240 PubMed abstracts. These datasets from a complex data environment for cognitive tasks on arsenic in foods including decision making, sense making and learning.

The datasets were the data sources for conducting visual analytic tasks to test the research hypothesis and accomplish the research objectives. Several key insights were made in this doctoral research. The concentration of inorganic arsenic in instant rice was lower than those of other rice types. This could be attributed to the pre-cooking and dehydration prior to reaching consumers. The concentration of Dimethylarsinic Acid (DMA) in a wild rice (aquatic grass) was notably lower than rice varieties (e.g. 0.0099 ppm versus 0.182 for a long grain white rice). This lower accumulation of DMA in rice grain could be due to differences in translocation of arsenic by *Oryza* species (true rice) and aquatic grasses. A view that grouped rice samples by country of origin and total arsenic concentration prioritized Malaysia for further insights on recently determined estimates of arsenic in rice sold or grown in Malaysia. The elemental arsenic concentrations determined by instrumental neutron activation analysis for 12 types of rice were obtained for literature. Additionally, daily intake values for arsenic via white rice for 7 Asian countries including Malaysia (33 µg) was retrieved from literature.

The 328 table ready foods were categorized into 12 food categories (baby foods; beverage; dairy; eggs; fats/oils; fruits; grains; legumes; meat, poultry, fish; mixtures; sweet and vegetables. This categorization of the table ready foods is a contribution of this research and enhances the communication of the TDS data available on the United States Food and Drug Administration (US FDA) website. Furthermore, views constructed provided integration of data from four total diet studies leading to identification of outlier concentrations for rice.

The 193 rice samples analysed by U.S. FDA were grouped into two groups using a cut-off level of 3 mcg of inorganic arsenic per serving. The visual analytics view constructed allows users to specify cut-off levels desired. A total of 86 sentences from 53 PubMed abstracts were identified as informative for arsenic concentrations. The sentences enabled literature curation for arsenic concentration and additional supporting information such as location of the research. An informative sentence provided global “normal” range of 0.08 to 0.20 mg/kg for arsenic in rice. A visual analytics resource developed was a dashboard that facilitates the interaction with text and a connection to the knowledge base of the PubMed literature database.

The presence of arsenic as a contaminant in foods continues to receive the attention of food and health safety authorities worldwide. Thus there are numerous research investigation and monitoring activities to (i) guide setting limits for arsenic in foods and (ii) provide best practices to reduce or eliminate arsenic in foods.

6.2 Recommendations

The research reported provides a foundation for additional investigations on visual analytics of data on arsenic concentrations in foods. In 2010, an article was published from contents of this thesis (Appendix 6). The article has been cited 9 times according to Google Scholar (http://scholar.google.com). These articles include those on visual analytics; rice, arsenic, and bacteria. Examples of environmental science related journals where the citing articles were published are: Ecotoxicology, Environmental Forensics, and Geosciences Journal. The article was one of the selections by the publishers in Chapter 2 (Environmental Health) of a book titled “Issues in Environment, Health, and Pollution: 2011 Edition” (Appendix 6). This evidence supports the importance of the research conducted and dissemination has resulted in inspiring further research.

Considering the multivariate and heterogeneous nature of data associated with contaminants in foods, the development of visual analytics tools are needed to facilitate diverse human cognitive tasks. Visual analytics tools can provide integrated automated analysis, interaction with data and data visualization critically needed to enhance decision making. The multivariate and heterogeneous nature of the datasets on arsenic in foods will require multivariate analysis using visual analytics software.

Stakeholders that would benefit include consumers; food and health safety personnel; farmers; and food producers. Arsenic content of baby foods warrants attention because of the early life exposures that could have life time adverse health consequences.

The actions of microorganisms in the soil are associated with availability of arsenic species for uptake by plants. Genomic data on microbial communities presents wealth of data for identifying mitigation strategies for arsenic uptake by plants. Arsenic metabolism pathways encoded in microbial genomes warrants further research. Visual analytics tasks could facilitate the discovery of biological processes for mitigating arsenic uptake from soil.

The increasing availability of central resources on data from total diet studies and research investigations presents a need for personnel with diverse levels of skills in data management and analysis. Training workshops and courses on the foundations and applications of visual analytics can contribute to global workforce development in food safety and environmental health. Research investigations could determine learning gains accomplished through hardware and software for visual analytics.

Finally, there is need to develop and evaluate informatics tools that have visual analytics capabilities in the domain of contaminants in foods.

REFERENCES

ACKERMAN, A. H., CREED, P. A., PARKS, A. N., FRICKE, M. W., SCHWEGEL, C. A., CREED, J. T., HEITKEMPER, D. T. & VELAL, N. P. 2005. Comparison of a chemical and enzymatic extraction of arsenic from rice and an assessment of the arsenic absorption from contaminated water by cooked rice. *Environ Sci Technol,* 39**,** 5241-6.

AGRE, P. & KOZONO, D. 2003. Aquaporin water channels: molecular mechanisms for human diseases. *FEBS Lett,* 555**,** 72-8.

AL-HAJJ, S., PIKE, I. & FISHER, B. 2013. Interactive Dashboards: Using Visual Analytics for knowledge Transfer and Decision Support. *Public Health's Wicked Problems: Can InfoVis Save Lives? IEEE Visual Analytics Science and Technology, Information Visualization, Scientific Visualization.***,** 1-8.

AL RMALLI, S. W., HARIS, P. I., HARRINGTON, C. F. & AYUB, M. 2005. A survey of arsenic in foodstuffs on sale in the United Kingdom and imported from Bangladesh. *Sci Total Environ,* 337**,** 23-30.

ALEX, B., GROVER, C., HADDOW, B., KABADJOR, M., KLEIN, E., MATTHEWS, M., ROEBUCK, S., TOBIN, R. & WANG, X. Assisted Curation: Does Text Mining Really Help? Pacific Symposium on Biocomputing, 2008. Citeseer, 556-567.

ALMELA, C., ALGORA, S., BENITO, V., CLEMENTE, M. J., DEVESA, V., SUNER, M. A., VELEZ, D. & MONTORO, R. 2002. Heavy metal, total arsenic, and inorganic arsenic contents of algae food products. *J Agric Food Chem,* 50**,** 918-23.

ANDRIENKO, G., ANDRIENKO, N. & WROBEL, S. 2007. Visual analytics tools for analysis of movement data. *ACM SIGKDD Explorations Newsletter,* 9**,** 38-46.

ANZFA 2001. The 19th Australian total diet survey. *Australia & New Zealand Food Authority: ANZFA*.

ARAGON, C. R., POON, S. S., ALDERING, G. S., THOMAS, R. C. & QUIMBY, R. Using visual analytics to maintain situation awareness in astrophysics. Visual Analytics Science and Technology, 2008. VAST'08. IEEE Symposium on, 2008. IEEE, 27-34.

BAASTRUP, R., SORENSEN, M., BALSTROM, T., FREDERIKSEN, K., LARSEN, C. L., TJONNELAND, A., OVERVAD, K. & RAASCHOU-NIELSEN, O. 2008. Arsenic in drinking-water and risk for cancer in Denmark. *Environ Health Perspect,* 116**,** 231-7.

BANERJEE, M., BANERJEE, N., BHATTACHARJEE, P., MONDAL, D., LYTHGOE, P. R., MARTINEZ, M., PAN, J., POLYA, D. A. & GIRI, A. K. 2013. High arsenic in rice is associated with elevated genotoxic effects in humans. *Sci Rep,* 3**,** 2195.

BEAMER, P., KEY, M. E., FERGUSON, A. C., CANALES, R. A., AUYEUNG, W. & LECKIE, J. O. 2008. Quantified activity pattern data from 6 to 27-month-old farmworker children for use in exposure assessment. *Environ Res,* 108**,** 239-46.

BENRAMDANE, L., ACCOMINOTTI, M., FANTON, L., MALICIER, D. & VALLON, J. J. 1999. Arsenic speciation in human organs following fatal arsenic trioxide poisoning--a case report. *Clin Chem,* 45**,** 301-6.

BETSY, A., SUDERSHAN RAO, V. & POLASA, K. 2012. Evolution of approaches in conducting total diet studies. *J Appl Toxicol,* 32**,** 765-76.

BORAK, J. & HOSGOOD, H. D. 2007. Seafood arsenic: implications for human risk assessment. *Regul Toxicol Pharmacol,* 47**,** 204-12.

BOWDEN, E. M., JUNG-BEEMAN, M., FLECK, J. & KOUNIOS, J. 2005. New approaches to demystifying insight. *Trends Cogn Sci,* 9**,** 322-8.
BRAMMER, H. & RAVENSCROFT, P. 2009. Arsenic in groundwater: a threat to sustainable agriculture in South and South-east Asia. *Environ Int,* 35**,** 647-54.
BROEKSEMA, B., BAUDEL, T., TELEA, A. G. & CRISAFULLI, P. 2013. Decision exploration lab: a visual analytics solution for decision management. *IEEE Trans Vis Comput Graph,* 19**,** 1972-81.
BUCHET, J. P. & LAUWERYS, R. 1985. Study of inorganic arsenic methylation by rat liver in vitro: relevance for the interpretation of observations in man. *Arch Toxicol,* 57**,** 125-9.
BURGER, J. & GOCHFELD, M. 2013. Selenium/mercury molar ratios in freshwater, marine, and commercial fish from the USA: variation, risk, and health management. *Rev Environ Health,* 28**,** 129-43.
BURLEY, D. & ASHBURN, V. 2010. Information Visualization As A Knowledge Integration Tool. *Journal of Knowledge Management Practice,* 11.
CADWALADER, G. O., RENSHAW, C. E., JACKSON, B. P., MAGILLIGAN, F. J., LANDIS, J. D. & BOSTICK, B. C. 2011. Erosion and physical transport via overland flow of arsenic and lead bound to silt-sized particles. *Geomorphology,* 128**,** 85-91.
CAMON, E. B., BARRELL, D. G., DIMMER, E. C., LEE, V., MAGRANE, M., MASLEN, J., BINNS, D. & APWEILER, R. 2005. An evaluation of GO annotation retrieval for BioCreAtIvE and GOA. *BMC Bioinformatics,* 6 Suppl 1**,** S17.
CARBREY, J. M., GORELICK-FELDMAN, D. A., KOZONO, D., PRAETORIUS, J., NIELSEN, S. & AGRE, P. 2003. Aquaglyceroporin AQP9: solute permeation and metabolic control of expression in liver. *Proc Natl Acad Sci U S A,* 100**,** 2945-50.
CARD, S. K., MACKINLAY, J. D. & SCHNEIDERMAN, B. 1999. *Readings in information visualization: using vision to think*, Morgan Kaufmann.
CAREY, A. M., SCHECKEL, K. G., LOMBI, E., NEWVILLE, M., CHOI, Y., NORTON, G. J., CHARNOCK, J. M., FELDMANN, J., PRICE, A. H. & MEHARG, A. A. 2010. Grain unloading of arsenic species in rice. *Plant Physiol,* 152**,** 309-19.
CHABOT, C. 2009. Demystifying visual analytics. *IEEE Comput Graph Appl,* 29**,** 84-7.
CHANG, R., ZIEMKIEWICZ, C., GREEN, T. M. & RIBARSKY, W. 2009. Defining insight for visual analytics. *Computer Graphics and Applications, IEEE,* 29**,** 14-17.
CHEN, C. & CZERWINSKI, M. P. 2000. Empirical evaluation of information visualizations: an introduction. *International Journal of Human-Computer Studies,* 53**,** 631-635.
CHEN, J. & GAO, J. 1993. The Chinese total diet study in 1990. Part II. Nutrients. *J AOAC Int,* 76**,** 1206-13.
CHEN, Y., WANG, Y., DUAN, X., ZHANG, H. & JIA, Y. 2014. Text Mining for Information Screen in Risk Assessment of Environmental Endocrine Disruptive Chemicals. *Advanced in Computer Science and its Applications.* Springer.
CHOI, M., YEPES, A. J., ZOBEL, J. & VERSPOOR, K. NEROC: Named Entity Recognizer of Chemicals. BioCreative Challenge Evaluation Workshop vol. 2, 2013. 97.
CLARKE, S., DE-GUSSEM, K. & BARNES, J. 2003. Flagellated protozoan infections in turkeys. *World Poultry,* 19**,** 1-4.

COBB, G. P., SANDS, K., WATERS, M., WIXSON, B. G. & DORWARD-KING, E. 2000. Accumulation of heavy metals by vegetables grown in mine wastes. *Environmental Toxicology and Chemistry,* 19**,** 600-607.

CODEX 2012. Proposed Draft Maximum Levels of Arsenic in Rice (At Step 3).

CONCHA, G., NERMELL, B. & VAHTER, M. V. 1998. Metabolism of inorganic arsenic in children with chronic high arsenic exposure in northern Argentina. *Environ Health Perspect,* 106**,** 355-9.

CONSUMER REPORTS 2012. Arsenic in your juice: how much is too much? Federal limits don't exist. *Consum Rep,* 77**,** 22-7.

CONSUMER REPORTS 2013. FDA aims to limit arsenic in apple juice. *Consum Rep,* 78**,** 15.

D'AMATO, M., TURRINI, A., AURELI, F., MORACCI, G., RAGGI, A., CHIARAVALLE, E., MANGIACOTTI, M., CENCI, T., ORLETTI, R., CANDELA, L., DI SANDRO, A. & CUBADDA, F. 2013. Dietary exposure to trace elements and radionuclides: the methodology of the Italian Total Diet Study 2012-2014. *Ann Ist Super Sanita,* 49**,** 272-80.

DAS, H. K., MITRA, A. K., SENGUPTA, P. K., HOSSAIN, A., ISLAM, F. & RABBANI, G. H. 2004. Arsenic concentrations in rice, vegetables, and fish in Bangladesh: a preliminary study. *Environ Int,* 30**,** 383-7.

DAUPHINE, D. C., FERRECCIO, C., GUNTUR, S., YUAN, Y., HAMMOND, S. K., BALMES, J., SMITH, A. H. & STEINMAUS, C. 2011. Lung function in adults following in utero and childhood exposure to arsenic in drinking water: preliminary findings. *Int Arch Occup Environ Health,* 84**,** 591-600.

DAVIS, A. P., MURPHY, C. G., JOHNSON, R., LAY, J. M., LENNON-HOPKINS, K., SARACENI-RICHARDS, C., SCIAKY, D., KING, B. L., ROSENSTEIN, M. C., WIEGERS, T. C. & MATTINGLY, C. J. 2013. The Comparative Toxicogenomics Database: update 2013. *Nucleic Acids Res,* 41**,** D1104-14.

DEL RAZO, L. M., GARCIA-VARGAS, G. G., VARGAS, H., ALBORES, A., GONSEBATT, M. E., MONTERO, R., OSTROSKY-WEGMAN, P., KELSH, M. & CEBRIAN, M. E. 1997. Altered profile of urinary arsenic metabolites in adults with chronic arsenicism. A pilot study. *Arch Toxicol,* 71**,** 211-7.

DELNOMDEDIEU, M., STYBLO, M. & THOMAS, D. J. 1995. Time dependence of accumulation and binding of inorganic and organic arsenic species in rabbit erythrocytes. *Chem Biol Interact,* 98**,** 69-83.

DONG, Z., WANG, H., DONG, Y., WANG, Y., LIU, W., MIAO, G., LIN, X., WANG, D. & LIU, B. 2013. Extensive microsatellite variation in rice induced by introgression from wild rice (Zizania latifolia Griseb.). *PLoS One,* 8**,** e62317.

DOPP, E., VON RECKLINGHAUSEN, U., DIAZ-BONE, R., HIRNER, A. V. & RETTENMEIER, A. W. 2010. Cellular uptake, subcellular distribution and toxicity of arsenic compounds in methylating and non-methylating cells. *Environ Res,* 110**,** 435-42.

DROBNA, Z., STYBLO, M. & THOMAS, D. J. 2009. In vitro assays of inorganic arsenic methylation. *Curr Protoc Toxicol,* 42**,** 4 32 1-4 32 10.

DROBNA, Z., WALTON, F. S., PAUL, D. S., XING, W., THOMAS, D. J. & STYBLO, M. 2010. Metabolism of arsenic in human liver: the role of membrane transporters. *Arch Toxicol,* 84**,** 3-16.

DUKER, A. A., CARRANZA, E. J. & HALE, M. 2005. Arsenic geochemistry and health. *Environ Int,* 31**,** 631-41.

EGAN, K. 2013. United States Food and Drug Administration’s Total Diet Study Program. *Total Diet Studies.* Springer.

EGAN, S. K., BOLGER, P. M. & CARRINGTON, C. D. 2007. Update of US FDA's Total Diet Study food list and diets. *J Expo Sci Environ Epidemiol,* 17**,** 573-82.
EGAN, S. K., TAO, S. S., PENNINGTON, J. A. & BOLGER, P. M. 2002. US Food and Drug Administration's Total Diet Study: intake of nutritional and toxic elements, 1991-96. *Food Addit Contam,* 19**,** 103-25.
EUROPEAN FOOD SAFETY AUTHORITY 2009. Panel on contaminants in the food chain (CONTAM); scientific opinion on arsenic in food. *EFSA J,* 7.
FAITA, F., CORI, L., BIANCHI, F. & ANDREASSI, M. G. 2013. Arsenic-induced genotoxicity and genetic susceptibility to arsenic-related pathologies. *Int J Environ Res Public Health,* 10**,** 1527-46.
FLYNN, C. A. 2013. Communicating Results in a Total Diet Study. *Total Diet Studies.* Springer.
FSAI 2011. Report on a Total Diet Study Carried out by the Food Safety Authority of Ireland in the Period 2001-2005. *Monitoring and Surviellance Series Chemical.* Dublin: Food Safety Authority of Ireland.
GAMBLE, M. V., LIU, X., AHSAN, H., PILSNER, J. R., ILIEVSKI, V., SLAVKOVICH, V., PARVEZ, F., CHEN, Y., LEVY, D., FACTOR-LITVAK, P. & GRAZIANO, J. H. 2006. Folate and arsenic metabolism: a double-blind, placebo-controlled folic acid-supplementation trial in Bangladesh. *Am J Clin Nutr,* 84**,** 1093-101.
GARCIA-ESQUINAS, E., POLLAN, M., UMANS, J. G., FRANCESCONI, K. A., GOESSLER, W., GUALLAR, E., HOWARD, B., FARLEY, J., BEST, L. G. & NAVAS-ACIEN, A. 2013. Arsenic exposure and cancer mortality in a US-based prospective cohort: the strong heart study. *Cancer Epidemiol Biomarkers Prev,* 22**,** 1944-53.
GARDNER, R. M., ENGSTROM, K., BOTTAI, M., HOQUE, W. A., RAQIB, R., BROBERG, K. & VAHTER, M. 2012. Pregnancy and the methyltransferase genotype independently influence the arsenic methylation phenotype. *Pharmacogenet Genomics,* 22**,** 508-16.
GARDNER, R. M., NERMELL, B., KIPPLER, M., GRANDER, M., LI, L., EKSTROM, E. C., RAHMAN, A., LONNERDAL, B., HOQUE, A. M. & VAHTER, M. 2011. Arsenic methylation efficiency increases during the first trimester of pregnancy independent of folate status. *Reprod Toxicol,* 31**,** 210-8.
GARELICK, H., JONES, H., DYBOWSKA, A. & VALSAMI-JONES, E. 2008. Arsenic pollution sources. *Rev Environ Contam Toxicol,* 197**,** 17-60.
GHOSH, A., AWAL, M. A., MAJUMDER, S., MOSTOFA, M., KHAIR, A., ISLAM, M. Z. & RAO, D. R. 2012. Arsenic in eggs and excreta of laying hens in Bangladesh: a preliminary study. *J Health Popul Nutr,* 30**,** 383-93.
GILBERT-DIAMOND, D., COTTINGHAM, K. L., GRUBER, J. F., PUNSHON, T., SAYARATH, V., GANDOLFI, A. J., BAKER, E. R., JACKSON, B. P., FOLT, C. L. & KARAGAS, M. R. 2011. Rice consumption contributes to arsenic exposure in US women. *Proc Natl Acad Sci U S A,* 108**,** 20656-60.
GIMOU, M. M., CHARRONDIERE, U. R., LEBLANC, J. C., NOEL, L., GUERIN, T. & POUILLOT, R. 2013. Dietary exposure and health risk assessment for 11 minerals and trace elements in Yaounde: the Cameroonian Total Diet Study. *Food Addit Contam Part A Chem Anal Control Expo Risk Assess,* 30**,** 1556-72.
GIMOU, M. M., CHARRONDIERE, U. R., LEBLANC, J. C. & POUILLOT, R. 2008. Dietary exposure to pesticide residues in Yaounde: the Cameroonian total diet study. *Food Addit Contam Part A Chem Anal Control Expo Risk Assess,* 25**,** 458-71.

GOMEZ-RUBIO, P., KLIMENTIDIS, Y. C., CANTU-SOTO, E., MEZA-MONTENEGRO, M. M., BILLHEIMER, D., LU, Z., CHEN, Z. & KLIMECKI, W. T. 2012. Indigenous American ancestry is associated with arsenic methylation efficiency in an admixed population of northwest Mexico. *J Toxicol Environ Health A,* 75**,** 36-49.

GUNDERSON, E. L. 1995. FDA Total Diet Study, July 1986-April 1991, dietary intakes of pesticides, selected elements, and other chemicals. *J AOAC Int,* 78**,** 1353-63.

HALDER, D., BHOWMICK, S., BISWAS, A., CHATTERJEE, D., NRIAGU, J., GUHA MAZUMDER, D. N., SLEJKOVEC, Z., JACKS, G. & BHATTACHARYA, P. 2013. Risk of arsenic exposure from drinking water and dietary components: implications for risk management in rural Bengal. *Environ Sci Technol,* 47**,** 1120-7.

HALL, M. N., NIEDZWIECKI, M., LIU, X., HARPER, K. N., ALAM, S., SLAVKOVICH, V., ILIEVSKI, V., LEVY, D., SIDDIQUE, A. B., PARVEZ, F., MEY, J. L., VAN GEEN, A., GRAZIANO, J. & GAMBLE, M. V. 2013. Chronic arsenic exposure and blood glutathione and glutathione disulfide concentrations in bangladeshi adults. *Environ Health Perspect,* 121**,** 1068-74.

HARGER, J. R. & CROSSNO, P. J. Comparison of open-source visual analytics toolkits. IS&T/SPIE Electronic Imaging, 2012. International Society for Optics and Photonics, 82940E-82940E-10.

HEGARTY, M. 2011. The Cognitive Science of Visual-Spatial Displays: Implications for Design. *Topics in cognitive science,* 3**,** 446-474.

HEITKEMPER, D. T., VELA, N. P., STEWART, K. R. & WESTPHAL, C. S. 2001. Determination of total and speciated arsenic in rice by ion chromatography and inductively coupled plasma mass spectrometry. *Journal of Analytical Atomic Spectrometry,* 16**,** 299-306.

HITE, A. H. 2013. Arsenic and rice: a call for regulation. *Nutrition,* 29**,** 353-4.

HOPENHAYN-RICH, C., BIGGS, M. L., SMITH, A. H., KALMAN, D. A. & MOORE, L. E. 1996. Methylation study of a population environmentally exposed to arsenic in drinking water. *Environ Health Perspect,* 104**,** 620-8.

HSUEH, L. 2013. Beyond regulations: Industry voluntary ban in arsenic use. *Journal of environmental management,* 131**,** 435-446.

HUGHES, M. F. 2006. Biomarkers of exposure: a case study with inorganic arsenic. *Environ Health Perspect,* 114**,** 1790-6.

HUGHES, M. F. 2009. Arsenic methylation, oxidative stress and cancer--is there a link? *J Natl Cancer Inst,* 101**,** 1660-1.

ISOKPEHI, R. D., COHLY, H. H., ANYANWU, M. N., RAJNARAYANAN, R. V., TCHOUNWOU, P. B., UDENSI, U. K. & GRAHAM-EVANS, B. E. 2010. Candidate single nucleotide polymorphism markers for arsenic responsiveness of protein targets. *Bioinform Biol Insights,* 4**,** 99-111.

ISOKPEHI, R. D., UDENSI, U. K., ANYANWU, M. N., MBAH, A. N., JOHNSON, M. O., EDUSEI, K., BAUER, M. A., HALL, R. A. & AWOFOLU, O. R. 2012. Knowledge building insights on biomarkers of arsenic toxicity to keratinocytes and melanocytes. *Biomark Insights,* 7**,** 127-41.

JACKSON, B. P., TAYLOR, V. F., KARAGAS, M. R., PUNSHON, T. & COTTINGHAM, K. L. 2012. Arsenic, organic foods, and brown rice syrup. *Environ Health Perspect,* 120**,** 623-6.

JAYASUMANA, M., PARANAGAMA, P., AMARASINGHE, M., WIJEWARDANE, K., DAHANAYAKE, K., FONSEKA, S., RAJAKARUNA, K., MAHAMITHAWA, A., SAMARASINGHE, U. & SENANAYAKE, V. 2013. Possible link of Chronic

arsenic toxicity with Chronic Kidney Disease of unknown etiology in Sri Lanka. *Journal of Natural Sciences Research,* 3**,** 64-73.
JAYATILAKE, N., MENDIS, S., MAHEEPALA, P., MEHTA, F. R. & TEAM, C. K. N. R. P. 2013. Chronic kidney disease of uncertain aetiology: prevalence and causative factors in a developing country. *BMC Nephrol,* 14**,** 180.
JEFCA 1989. Evaluation of certain food additives and contaminants: thirty-third report of the Joint FAO/WHO Expert Committee on Food Additives [meeting held in Geneva from 21 to 30 March 1988].
JEFCA 1995. Evaluation of certain food additives and contaminants: forty-fourth report of the Joint FAO/WHO Expert Committee on Food Additives.
JEFCA 2011. Safety evaluation of certain contaminants in food: prepared by the Seventy-second meeting of the Joint FAO/WHO Expert Committee on Food Additives (JECFA).
JIA, Y., HUANG, H., SUN, G. X., ZHAO, F. J. & ZHU, Y. G. 2012. Pathways and relative contributions to arsenic volatilization from rice plants and paddy soil. *Environ Sci Technol,* 46**,** 8090-6.
JIA, Y., HUANG, H., ZHONG, M., WANG, F. H., ZHANG, L. M. & ZHU, Y. G. 2013. Microbial arsenic methylation in soil and rice rhizosphere. *Environ Sci Technol,* 47**,** 3141-8.
JOLLIFFE, D. M. 1993. A history of the use of arsenicals in man. *J R Soc Med,* 86**,** 287-9.
KAWALEK, J. 2011. Final Report on Study 275.30. Provide Data on Various Arsenic Species Present in Broilers Treated with Roxarsone: Comparison with Untreated Birds. . Laurel, Rockville, and College Park, MD: Center for Veterinary Medicine/Office of Regulatory Science, : US Food and Drug Administration
KEIM, D., ANDRIENKO, G., FEKETE, J.-D., GÖRG, C., KOHLHAMMER, J. & MELANÇON, G. 2008. *Visual analytics: Definition, process, and challenges*, Springer.
KEIM, D. A., KOHLHAMMER, J., ELLIS, G. & MANSMANN, F. 2010. *Mastering The Information Age-Solving Problems with Visual Analytics*, Florian Mansmann.
KHAN, N. I., BRUCE, D., NAIDU, R. & OWENS, G. 2009. Implementation of food frequency questionnaire for the assessment of total dietary arsenic intake in Bangladesh: part B, preliminary findings. *Environ Geochem Health,* 31 Suppl 1**,** 221-38.
KIM, J. D., OHTA, T. & TSUJII, J. 2008. Corpus annotation for mining biomedical events from literature. *BMC Bioinformatics,* 9**,** 10.
KLASSEN, R., DOUMA, S. & RENCZ, A. 2010. Environmental and Human Health Risk Assessment for Essential Trace Elements: Considering the Role for Geoscience. *Journal of Toxicology and Environmental Health, Part A,* 73**,** 242-252.
KOBSA, A. An empirical comparison of three commercial information visualization systems. Information Visualization, 2001. INFOVIS 2001. IEEE Symposium on, 2001. IEEE, 123-130.
KWON, E., ZHANG, H., WANG, Z., JHANGRI, G. S., LU, X., FOK, N., GABOS, S., LI, X. F. & LE, X. C. 2004. Arsenic on the hands of children after playing in playgrounds. *Environ Health Perspect,* 112**,** 1375-80.
LASKY, T. 2013. Arsenic levels in chicken. *Environ Health Perspect,* 121**,** A267.

LASKY, T., SUN, W., KADRY, A. & HOFFMAN, M. K. 2004. Mean total arsenic concentrations in chicken 1989-2000 and estimated exposures for consumers of chicken. *Environ Health Perspect,* 112**,** 18-21.
LEE, H. J., SHIM, S. H., SONG, M. R., LEE, H. & PARK, J. C. 2013. CoMAGC: a corpus with multi-faceted annotations of gene-cancer relations. *BMC Bioinformatics,* 14**,** 323.
LEHRER, J. 2008. The eureka hunt. *New Yorker*.
LIU, W.-J., MCGRATH, S. P. & ZHAO, F.-J. 2013. Silicon has opposite effects on the accumulation of inorganic and methylated arsenic species in rice. *Plant and Soil***,** 1-9.
LIU, W. J., ZHU, Y. G., SMITH, F. A. & SMITH, S. E. 2004. Do phosphorus nutrition and iron plaque alter arsenate (As) uptake by rice seedlings in hydroponic culture? *New Phytologist,* 162**,** 481-488.
LIU, Z., SHEN, J., CARBREY, J. M., MUKHOPADHYAY, R., AGRE, P. & ROSEN, B. P. 2002. Arsenite transport by mammalian aquaglyceroporins AQP7 and AQP9. *Proc Natl Acad Sci U S A,* 99**,** 6053-8.
LIU, Z., STYBLO, M. & ROSEN, B. P. 2006. Methylarsonous acid transport by aquaglyceroporins. *Environ Health Perspect,* 114**,** 527-31.
LOMAX, C., LIU, W. J., WU, L., XUE, K., XIONG, J., ZHOU, J., MCGRATH, S. P., MEHARG, A. A., MILLER, A. J. & ZHAO, F. J. 2012. Methylated arsenic species in plants originate from soil microorganisms. *New Phytol,* 193**,** 665-72.
LOMBI, E., SCHECKEL, K. G., PALLON, J., CAREY, A. M., ZHU, Y. G. & MEHARG, A. A. 2009. Speciation and distribution of arsenic and localization of nutrients in rice grains. *New Phytol,* 184**,** 193-201.
LUH, B. S. 1991. Quick-cooking rice. *Rice Utilizafion,* 11**,** 121-146.
MACHILIS, S. 2013. *Chart and image gallery: 30+ free tools for data visualization and analysis* [Online]. Available: http://www.computerworld.com/s/article/9214755/Chart_and_image_gallery_30_free_tools_for_data_visualization_and_analysis.
MACKINLAY, J., HANRAHAN, P. & STOLTE, C. 2007. Show me: automatic presentation for visual analysis. *IEEE Trans Vis Comput Graph,* 13**,** 1137-44.
MAHER, W., FOSTER, S., KRIKOWA, F., DONNER, E. & LOMBI, E. 2013. Measurement of inorganic arsenic species in rice after nitric acid extraction by HPLC-ICPMS: verification using XANES. *Environ Sci Technol,* 47**,** 5821-7.
MAHIMAIRAJA, S., BOLAN, N. S., ADRIANO, D. & ROBINSON, B. 2005. Arsenic contamination and its risk management in complex environmental settings. *Advances in Agronomy,* 86**,** 1-82.
MAI, X. Q., LUO, J., WU, J. H. & LUO, Y. J. 2004. “Aha!” effects in a guessing riddle task: An event-related potential study. *Human brain mapping,* 22**,** 261-270.
MARIN, A., MASSCHELEYN, P. & PATRICK JR, W. 1992. The influence of chemical form and concentration of arsenic on rice growth and tissue arsenic concentration. *Plant and Soil,* 139**,** 175-183.
MARTINEZ, V. D., VUCIC, E. A., BECKER-SANTOS, D. D., GIL, L. & LAM, W. L. 2011. Arsenic exposure and the induction of human cancers. *J Toxicol,* 2011**,** 431287.
MASS, M. J., TENNANT, A., ROOP, B. C., CULLEN, W. R., STYBLO, M., THOMAS, D. J. & KLIGERMAN, A. D. 2001. Methylated trivalent arsenic species are genotoxic. *Chem Res Toxicol,* 14**,** 355-61.
MATSCHULLAT, J. 2000. Arsenic in the geosphere--a review. *Sci Total Environ,* 249**,** 297-312.

MBAH, A. N., MAHMUD, O., AWOFOLU, O. R. & ISOKPEHI, R. D. 2013. Inferences on the biochemical and environmental regulation of universal stress proteins from Schistosomiasis parasites. *Adv Appl Bioinform Chem,* 6**,** 15-27.

MCDERMOTT, J. R., JIANG, X., BEENE, L. C., ROSEN, B. P. & LIU, Z. 2010. Pentavalent methylated arsenicals are substrates of human AQP9. *Biometals,* 23**,** 119-27.

MEHARG, A. A. 2004. Arsenic in rice--understanding a new disaster for South-East Asia. *Trends Plant Sci,* 9**,** 415-7.

MEHARG, A. A., WILLIAMS, P. N., ADOMAKO, E., LAWGALI, Y. Y., DEACON, C., VILLADA, A., CAMBELL, R. C., SUN, G., ZHU, Y. G., FELDMANN, J., RAAB, A., ZHAO, F. J., ISLAM, R., HOSSAIN, S. & YANAI, J. 2009. Geographical variation in total and inorganic arsenic content of polished (white) rice. *Environ Sci Technol,* 43**,** 1612-7.

MELKONIAN, S., ARGOS, M., HALL, M. N., CHEN, Y., PARVEZ, F., PIERCE, B., CAO, H., ASCHEBROOK-KILFOY, B., AHMED, A., ISLAM, T., SLAVCOVICH, V., GAMBLE, M., HARIS, P. I., GRAZIANO, J. H. & AHSAN, H. 2013. Urinary and dietary analysis of 18,470 bangladeshis reveal a correlation of rice consumption with arsenic exposure and toxicity. *PLoS One,* 8**,** e80691.

MENG, Y. L., LIU, Z. & ROSEN, B. P. 2004. As(III) and Sb(III) uptake by GlpF and efflux by ArsB in Escherichia coli. *J Biol Chem,* 279**,** 18334-41.

MERCER, T. & FROSTICK, L. 2012. Leaching characteristics of CCA-treated wood waste: A UK study. *Science of the Total Environment,* 427**,** 165-174.

MILLMAN, A., TANG, D. & PERERA, F. P. 2008. Air pollution threatens the health of children in China. *Pediatrics,* 122**,** 620-8.

MOON, J., EBIHARA, M., NI, B., ARPORN, B., SETYO, P., THERESIA, R., WEE, B., SALIM, N. A. A. & PABROA, P. C. 2012. A NAA collaborative study in white rice performed in seven Asian countries. *Journal of Radioanalytical and Nuclear Chemistry,* 291**,** 217-221.

MOON, K. A., GUALLAR, E., UMANS, J. G., DEVEREUX, R. B., BEST, L. G., FRANCESCONI, K. A., GOESSLER, W., POLLAK, J., SILBERGELD, E. K., HOWARD, B. V. & NAVAS-ACIEN, A. 2013. Association between exposure to low to moderate arsenic levels and incident cardiovascular disease. A prospective cohort study. *Ann Intern Med,* 159**,** 649-59.

MOONEY, C., ABBEY, J. L. & LAAJOKI, L. 2013. Managing Concentration Data—Validation, Security, and Interpretation. *Total Diet Studies.* Springer.

MOORE, K. L., SCHRODER, M., LOMBI, E., ZHAO, F. J., MCGRATH, S. P., HAWKESFORD, M. J., SHEWRY, P. R. & GROVENOR, C. R. 2010. NanoSIMS analysis of arsenic and selenium in cereal grain. *New Phytol,* 185**,** 434-45.

MOY, G. G. 2013. Total Diet Studies—What They Are and Why They Are Important. *Total Diet Studies.* Springer.

MOY, G. G. & VANNOORT, R. W. 2013. *Total Diet Studies*, Springer London, Limited.

MULLER, H. M., KENNY, E. E. & STERNBERG, P. W. 2004. Textpresso: an ontology-based information retrieval and extraction system for biological literature. *PLoS Biol,* 2**,** e309.

MUÑIZ ORTIZ, J. G., WALLACE, K. A., LEINISCH, F., KADIISKA, M. B., MASON, R. P. & KLIGERMAN, A. D. 2013. Catalase has a key role in protecting cells from the genotoxic effects of monomethylarsonous acid: A highly active metabolite of arsenic. *Environmental and molecular mutagenesis*.

MURRAY, D. 2013. *Tableau Your Data!: Fast and Easy Visual Analysis with Tableau Software*, John Wiley & Sons.

MUSIL, S., PETURSDOTTIR, A. H., RAAB, A., GUNNLAUGSDOTTIR, H., KRUPP, E. & FELDMANN, J. 2014. Speciation without Chromatography Using Selective Hydride Generation: Inorganic Arsenic in Rice and Samples of Marine Origin. *Anal Chem*.

NACHMAN, K. E., BARON, P. A., RABER, G., FRANCESCONI, K. A. & LOVE, D. C. 2013a. Arsenic levels in chicken: Nachman et al. respond. *Environ Health Perspect,* 121**,** A267-8.

NACHMAN, K. E., BARON, P. A., RABER, G., FRANCESCONI, K. A., NAVAS-ACIEN, A. & LOVE, D. C. 2013b. Roxarsone, inorganic arsenic, and other arsenic species in chicken: a U.S.-Based market basket sample. *Environ Health Perspect,* 121**,** 818-24.

NAUJOKAS, M. F., ANDERSON, B., AHSAN, H., APOSHIAN, H. V., GRAZIANO, J. H., THOMPSON, C. & SUK, W. A. 2013. The broad scope of health effects from chronic arsenic exposure: update on a worldwide public health problem. *Environ Health Perspect,* 121**,** 295-302.

NEI. 2012. *National Eye Institute Research Grant for Secondary Analysis (R21)* [Online]. Bethesda, MD: National Eye Institute. Available: http://grants.nih.gov/grants/guide/pa-files/PAR-13-035.html 2013].

NELSON, E. K., PIEHLER, B., RAUCH, A., RAMSAY, S., HOLMAN, D., ASARE, S., ASARE, A. & IGRA, M. 2013. Ancillary study management systems: a review of needs. *BMC Med Inform Decis Mak,* 13**,** 5.

OTLES, S. & CAGINDI, O. 2010. Health importance of arsenic in drinking water and food. *Environ Geochem Health,* 32**,** 367-71.

PAN, W. C., SEOW, W. J., KILE, M. L., HOFFMAN, E. B., QUAMRUZZAMAN, Q., RAHMAN, M., MAHIUDDIN, G., MOSTOFA, G., LU, Q. & CHRISTIANI, D. C. 2013. Association of low to moderate levels of arsenic exposure with risk of type 2 diabetes in Bangladesh. *Am J Epidemiol,* 178**,** 1563-70.

PETRICK, J. S., JAGADISH, B., MASH, E. A. & APOSHIAN, H. V. 2001. Monomethylarsonous acid (MMA(III)) and arsenite: LD(50) in hamsters and in vitro inhibition of pyruvate dehydrogenase. *Chem Res Toxicol,* 14**,** 651-6.

PIETROBON, R., GULLER, U., MARTINS, H., MENEZES, A. P., HIGGINS, L. D. & JACOBS, D. O. 2004. A suite of web applications to streamline the interdisciplinary collaboration in secondary data analyses. *BMC Med Res Methodol,* 4**,** 29.

PLAISANT, C., FEKETE, J.-D. & GRINSTEIN, G. 2008. Promoting insight-based evaluation of visualizations: From contest to benchmark repository. *Visualization and Computer Graphics, IEEE Transactions on,* 14**,** 120-134.

RADABAUGH, T. R. & APOSHIAN, H. V. 2000. Enzymatic reduction of arsenic compounds in mammalian systems: reduction of arsenate to arsenite by human liver arsenate reductase. *Chem Res Toxicol,* 13**,** 26-30.

RAHMAN, M. A. & HASEGAWA, H. 2011. High levels of inorganic arsenic in rice in areas where arsenic-contaminated water is used for irrigation and cooking. *Sci Total Environ,* 409**,** 4645-55.

RAHMAN, M. A., HASEGAWA, H., RAHMAN, M. A., RAHMAN, M. M. & MIAH, M. A. 2006. Influence of cooking method on arsenic retention in cooked rice related to dietary exposure. *Sci Total Environ,* 370**,** 51-60.

RAHMAN, M. A., HASEGAWA, H., RAHMAN, M. M., RAHMAN, M. A. & MIAH, M. A. 2007. Accumulation of arsenic in tissues of rice plant (Oryza sativa L.) and its distribution in fractions of rice grain. *Chemosphere,* 69**,** 942-8.

RAHMAN, M. M., OWENS, G. & NAIDU, R. 2009. Arsenic levels in rice grain and assessment of daily dietary intake of arsenic from rice in arsenic-contaminated regions of Bangladesh--implications to groundwater irrigation. *Environ Geochem Health,* 31 Suppl 1**,** 179-87.

RAMIREZ-SOLIS, A., MUKOPADHYAY, R., ROSEN, B. P. & STEMMLER, T. L. 2004. Experimental and theoretical characterization of arsenite in water: insights into the coordination environment of As-O. *Inorg Chem,* 43**,** 2954-9.

RAZO, I., CARRIZALES, L., CASTRO, J., DIAZ-BARRIGA, F. & MONROY, M. 2004. Arsenic and heavy metal pollution of soil, water and sediments in a semi-arid climate mining area in Mexico. *Water, Air, and Soil Pollution,* 152**,** 129-152.

RIEUWERTS, J. S., SEARLE, P. & BUCK, R. 2006. Bioaccessible arsenic in the home environment in southwest England. *Sci Total Environ,* 371**,** 89-98.

RINTALA, E. M., EKHOLM, P., KOIVISTO, P., PELTONEN, K. & VENALAINEN, E. R. 2014. The intake of inorganic arsenic from long grain rice and rice-based baby food in Finland - Low safety margin warrants follow up. *Food Chem,* 150**,** 199-205.

RODRIGUEZ-ESTEBAN, R. 2009. Biomedical text mining and its applications. *PLoS Comput Biol,* 5**,** e1000597.

ROSADO, J. L., RONQUILLO, D., KORDAS, K., ROJAS, O., ALATORRE, J., LOPEZ, P., GARCIA-VARGAS, G., DEL CARMEN CAAMANO, M., CEBRIAN, M. E. & STOLTZFUS, R. J. 2007. Arsenic exposure and cognitive performance in Mexican schoolchildren. *Environ Health Perspect,* 115**,** 1371-5.

ROSE, M., BAXTER, M., BRERETON, N. & BASKARAN, C. 2010. Dietary exposure to metals and other elements in the 2006 UK Total Diet Study and some trends over the last 30 years. *Food Addit Contam Part A Chem Anal Control Expo Risk Assess,* 27**,** 1380-404.

ROYCHOWDHURY, T., UCHINO, T., TOKUNAGA, H. & ANDO, M. 2002. Survey of arsenic in food composites from an arsenic-affected area of West Bengal, India. *Food Chem Toxicol,* 40**,** 1611-21.

SALAZAR, A. M., CALDERON-ARANDA, E., CEBRIAN, M. E., SORDO, M., BENDESKY, A., GOMEZ-MUNOZ, A., ACOSTA-SAAVEDRA, L. & OSTROSKY-WEGMAN, P. 2004. p53 expression in circulating lymphocytes of non-melanoma skin cancer patients from an arsenic contaminated region in Mexico. A pilot study. *Mol Cell Biochem,* 255**,** 25-31.

SALAZAR, A. M., OSTROSKY-WEGMAN, P., MENENDEZ, D., MIRANDA, E., GARCIA-CARRANCA, A. & ROJAS, E. 1997. Induction of p53 protein expression by sodium arsenite. *Mutat Res,* 381**,** 259-65.

SALIM, N. A. A., ELIAS, M. S., WOOD, A. K., SANURI, E., HAMZAH, M. S. & RAHMAN, S. A. 2010. MULTIELEMENT ANALYSIS IN RICE GRAINS BY INSTRUMENTAL NEUTRON ACTIVATION ANALYSIS. *Journal of Nuclear and Related Technologies,* 7.

SARAIYA, P., NORTH, C. & DUCA, K. 2005. An insight-based methodology for evaluating bioinformatics visualizations. *Visualization and Computer Graphics, IEEE Transactions on,* 11**,** 443-456.

SARAIYA, P., NORTH, C., LAM, V. & DUCA, K. A. 2006. An insight-based longitudinal study of visual analytics. *Visualization and Computer Graphics, IEEE Transactions on,* 12**,** 1511-1522.

SAVIKHIN, A., MACIEJEWSKI, R. & EBERT, D. S. Applied visual analytics for economic decision-making. Visual Analytics Science and Technology, 2008. VAST'08. IEEE Symposium on, 2008. IEEE, 107-114.

SCHMIDT, C. W. 2012. Maryland bans arsenical drug in chicken feed. *Environ Health Perspect,* 120**,** a269.
SCOTT, N., HATLELID, K. M., MACKENZIE, N. E. & CARTER, D. E. 1993. Reactions of arsenic(III) and arsenic(V) species with glutathione. *Chem Res Toxicol,* 6**,** 102-6.
SEDIG, K. & PARSONS, P. 2013. Interaction design for complex cognitive activities with visual representations: A pattern-based approach. *AIS Transactions on Human-Computer Interaction,* 5**,** 84-133.
SHEN, S., LI, X. F., CULLEN, W. R., WEINFELD, M. & LE, X. C. 2013. Arsenic binding to proteins. *Chem Rev,* 113**,** 7769-92.
SIMEONOVA, P. P. & LUSTER, M. I. 2000. Mechanisms of arsenic carcinogenicity: genetic or epigenetic mechanisms? *J Environ Pathol Toxicol Oncol,* 19**,** 281-6.
SIROT, V., GUERIN, T., VOLATIER, J. L. & LEBLANC, J. C. 2009. Dietary exposure and biomarkers of arsenic in consumers of fish and shellfish from France. *Sci Total Environ,* 407**,** 1875-85.
SLOTH, J. J. 2013. Arsenic speciation in food–current status on standardization of methods for specific determination of inorganic arsenic. *Annual NRL-EURL workshop EURL-CEFAO*.
SMITH, A. H., GOYCOLEA, M., HAQUE, R. & BIGGS, M. L. 1998. Marked increase in bladder and lung cancer mortality in a region of Northern Chile due to arsenic in drinking water. *Am J Epidemiol,* 147**,** 660-9.
SMITH, A. H., MARSHALL, G., LIAW, J., YUAN, Y., FERRECCIO, C. & STEINMAUS, C. 2012. Mortality in young adults following in utero and childhood exposure to arsenic in drinking water. *Environ Health Perspect,* 120**,** 1527-31.
SMITH, A. K., AYANIAN, J. Z., COVINSKY, K. E., LANDON, B. E., MCCARTHY, E. P., WEE, C. C. & STEINMAN, M. A. 2011. Conducting high-value secondary dataset analysis: an introductory guide and resources. *J Gen Intern Med,* 26**,** 920-9.
SMITH, D., RAO, R., LIUZZO, J. & CHAMPAGNE, E. 1985. Chemical Treatment and Process Modification for Producing Improved Quick-Cooking Rice. *Journal of Food science,* 50**,** 926-931.
SMITH, E., JUHASZ, A. L. & WEBER, J. 2009. Arsenic uptake and speciation in vegetables grown under greenhouse conditions. *Environ Geochem Health,* 31 Suppl 1**,** 125-32.
SMITH, E., NAIDU, R., WEBER, J. & JUHASZ, A. L. 2008. The impact of sequestration on the bioaccessibility of arsenic in long-term contaminated soils. *Chemosphere,* 71**,** 773-780.
SOIGNET, S. L., MASLAK, P., WANG, Z. G., JHANWAR, S., CALLEJA, E., DARDASHTI, L. J., CORSO, D., DEBLASIO, A., GABRILOVE, J., SCHEINBERG, D. A., PANDOLFI, P. P. & WARRELL, R. P., JR. 1998. Complete remission after treatment of acute promyelocytic leukemia with arsenic trioxide. *N Engl J Med,* 339**,** 1341-8.
SOMMERFELD, G. & MOY, G. G. 2013. OPAL—A Program to Manage Data on Chemicals in Food and the Diet. *Total Diet Studies.* Springer.
STAMATELOS, S. K., BRINKERHOFF, C. J., ISUKAPALLI, S. S. & GEORGOPOULOS, P. G. 2011. Mathematical model of uptake and metabolism of arsenic(III) in human hepatocytes - Incorporation of cellular antioxidant response and threshold-dependent behavior. *BMC Syst Biol,* 5**,** 16.
STRAIF, K., BENBRAHIM-TALLAA, L., BAAN, R., GROSSE, Y., SECRETAN, B., EL GHISSASSI, F., BOUVARD, V., GUHA, N., FREEMAN, C., GALICHET, L.,

COGLIANO, V. & GROUP, W. H. O. I. A. F. R. O. C. M. W. 2009. A review of human carcinogens--part C: metals, arsenic, dusts, and fibres. *Lancet Oncol,* 10**,** 453-4.
STYBLO, M., DEL RAZO, L. M., LECLUYSE, E. L., HAMILTON, G. A., WANG, C., CULLEN, W. R. & THOMAS, D. J. 1999. Metabolism of arsenic in primary cultures of human and rat hepatocytes. *Chem Res Toxicol,* 12**,** 560-5.
STYBLO, M., DELNOMDEDIEU, M. & THOMAS, D. J. 1996. Mono- and dimethylation of arsenic in rat liver cytosol in vitro. *Chem Biol Interact,* 99**,** 147-64.
STYBLO, M., DROBNA, Z., JASPERS, I., LIN, S. & THOMAS, D. J. 2002. The role of biomethylation in toxicity and carcinogenicity of arsenic: a research update. *Environ Health Perspect,* 110 Suppl 5**,** 767-71.
SUMI, D. & HIMENO, S. 2012. Role of arsenic (+3 oxidation state) methyltransferase in arsenic metabolism and toxicity. *Biol Pharm Bull,* 35**,** 1870-5.
SUMI, D., SHINKAI, Y. & KUMAGAI, Y. 2010. Signal transduction pathways and transcription factors triggered by arsenic trioxide in leukemia cells. *Toxicol Appl Pharmacol,* 244**,** 385-92.
SUN, G. X., WILLIAMS, P. N., CAREY, A. M., ZHU, Y. G., DEACON, C., RAAB, A., FELDMANN, J., ISLAM, R. M. & MEHARG, A. A. 2008. Inorganic arsenic in rice bran and its products are an order of magnitude higher than in bulk grain. *Environ Sci Technol,* 42**,** 7542-6.
TAKAHASHI, Y., MINAMIKAWA, R., HATTORI, K. H., KURISHIMA, K., KIHOU, N. & YUITA, K. 2004. Arsenic behavior in paddy fields during the cycle of flooded and non-flooded periods. *Environ Sci Technol,* 38**,** 1038-44.
TAO, S. S. & BOLGER, P. M. 1999. Dietary arsenic intakes in the United States: FDA Total Diet Study, September 1991-December 1996. *Food Addit Contam,* 16**,** 465-72.
TCHOUNWOU, P. B., PATLOLLA, A. K. & CENTENO, J. A. 2003. Carcinogenic and systemic health effects associated with arsenic exposure--a critical review. *Toxicol Pathol,* 31**,** 575-88.
THOMAS, D. J. 2007. Molecular processes in cellular arsenic metabolism. *Toxicol Appl Pharmacol,* 222**,** 365-73.
THOMAS, D. J., LI, J., WATERS, S. B., XING, W., ADAIR, B. M., DROBNA, Z., DEVESA, V. & STYBLO, M. 2007. Arsenic (+3 oxidation state) methyltransferase and the methylation of arsenicals. *Exp Biol Med (Maywood),* 232**,** 3-13.
THOMAS, D. J., STYBLO, M. & LIN, S. 2001. The cellular metabolism and systemic toxicity of arsenic. *Toxicol Appl Pharmacol,* 176**,** 127-44.
THOMAS, J. J. & COOK, K. A. 2005. *Illuminating the path: The research and development agenda for visual analytics*, IEEE Computer Society Press.
THOMAS, J. J. & COOK, K. A. 2006. A visual analytics agenda. *Computer Graphics and Applications, IEEE,* 26**,** 10-13.
TOKAR, E. J., KOJIMA, C. & WAALKES, M. P. 2013. Methylarsonous acid causes oxidative DNA damage in cells independent of the ability to biomethylate inorganic arsenic. *Archives of toxicology***,** 1-13.
TSENG, C. H. 2009. A review on environmental factors regulating arsenic methylation in humans. *Toxicol Appl Pharmacol,* 235**,** 338-50.
UNEYAMA, C., TODA, M., YAMAMOTO, M. & MORIKAWA, K. 2007. Arsenic in various foods: cumulative data. *Food Addit Contam,* 24**,** 447-534.
USATSDR 2007. Toxicological profile for arsenic. *Agency for Toxic Substances and Disease Registry, Division of Toxicology, Atlanta, GA*.

USFDA 2005. Beverages: bottled water. Final rule. *Fed Regist,* 70**,** 33694-701.
VAHTER, M. 2002. Mechanisms of arsenic biotransformation. *Toxicology,* 181-182**,** 211-7.
VAHTER, M. 2008. Health effects of early life exposure to arsenic. *Basic Clin Pharmacol Toxicol,* 102**,** 204-11.
VAHTER, M., CONCHA, G., NERMELL, B., NILSSON, R., DULOUT, F. & NATARAJAN, A. T. 1995. A unique metabolism of inorganic arsenic in native Andean women. *Eur J Pharmacol,* 293**,** 455-62.
VAHTER, M. & MARAFANTE, E. 1987. Effects of low dietary intake of methionine, choline or proteins on the biotransformation of arsenite in the rabbit. *Toxicol Lett,* 37**,** 41-6.
VINCENT, D., HASTINGS-TOLSMA, M. & EFFKEN, J. 2010. Data Visualization and Large Nursing Datasets. *Online Journal of Nursing Informatics (OJNI),* 14.
VOGT, B. L. & ROSSMAN, T. G. 2001. Effects of arsenite on p53, p21 and cyclin D expression in normal human fibroblasts -- a possible mechanism for arsenite's comutagenicity. *Mutat Res,* 478**,** 159-68.
VRACKO, P., TUOMISTO, J., GRAD, J. & KUNSELER, E. 2009. Exposure of children to chemical hazards in food.
WALTON, F. S., WATERS, S. B., JOLLEY, S. L., LECLUYSE, E. L., THOMAS, D. J. & STYBLO, M. 2003. Selenium compounds modulate the activity of recombinant rat AsIII-methyltransferase and the methylation of arsenite by rat and human hepatocytes. *Chem Res Toxicol,* 16**,** 261-5.
WANG, H., XI, S., LIU, Z., YANG, Y., ZHENG, Q., WANG, F., XU, Y., WANG, Y., ZHENG, Y. & SUN, G. 2013. Arsenic methylation metabolism and liver injury of acute promyelocytic leukemia patients undergoing arsenic trioxide treatment. *Environ Toxicol,* 28**,** 267-75.
WATERS, S. B., DEVESA, V., DEL RAZO, L. M., STYBLO, M. & THOMAS, D. J. 2004a. Endogenous reductants support the catalytic function of recombinant rat cyt19, an arsenic methyltransferase. *Chem Res Toxicol,* 17**,** 404-9.
WATERS, S. B., DEVESA, V., FRICKE, M. W., CREED, J. T., STYBLO, M. & THOMAS, D. J. 2004b. Glutathione modulates recombinant rat arsenic (+3 oxidation state) methyltransferase-catalyzed formation of trimethylarsine oxide and trimethylarsine. *Chem Res Toxicol,* 17**,** 1621-9.
WEI, Y., ZHU, J. & NGUYEN, A. 2013. Rice consumption and urinary concentrations of arsenic in US adults. *Int J Environ Health Res*.
WIEGERS, T. C., DAVIS, A. P., COHEN, K. B., HIRSCHMAN, L. & MATTINGLY, C. J. 2009. Text mining and manual curation of chemical-gene-disease networks for the comparative toxicogenomics database (CTD). *BMC Bioinformatics,* 10**,** 326.
WILBUR, W. J., RZHETSKY, A. & SHATKAY, H. 2006. New directions in biomedical text annotation: definitions, guidelines and corpus construction. *BMC Bioinformatics,* 7**,** 356.
WILLIAMS, P. N., RAAB, A., FELDMANN, J. & MEHARG, A. A. 2007a. Market basket survey shows elevated levels of As in South Central U.S. processed rice compared to California: consequences for human dietary exposure. *Environ Sci Technol,* 41**,** 2178-83.
WILLIAMS, P. N., VILLADA, A., DEACON, C., RAAB, A., FIGUEROLA, J., GREEN, A. J., FELDMANN, J. & MEHARG, A. A. 2007b. Greatly enhanced arsenic shoot assimilation in rice leads to elevated grain levels compared to wheat and barley. *Environ Sci Technol,* 41**,** 6854-9.

WINNENBURG, R., WACHTER, T., PLAKE, C., DOMS, A. & SCHROEDER, M. 2008. Facts from text: can text mining help to scale-up high-quality manual curation of gene products with ontologies? *Brief Bioinform,* 9**,** 466-78.
WONG, P. C. & THOMAS, J. 2004. Visual analytics. *IEEE Comput Graph Appl,* 24**,** 20-1.
WONG, W. W., CHUNG, S. W., CHAN, B. T., HO, Y. Y. & XIAO, Y. 2013. Dietary exposure to inorganic arsenic of the Hong Kong population: results of the first Hong Kong total diet study. *Food Chem Toxicol,* 51**,** 379-85.
XU, X. Y., MCGRATH, S. P., MEHARG, A. A. & ZHAO, F. J. 2008. Growing rice aerobically markedly decreases arsenic accumulation. *Environ Sci Technol,* 42**,** 5574-9.
XUE, J., ZARTARIAN, V., WANG, S. W., LIU, S. V. & GEORGOPOULOS, P. 2010. Probabilistic Modeling of Dietary Arsenic Exposure and Dose and Evaluation with 2003-2004 NHANES Data. *Environ Health Perspect,* 118**,** 345-50.
YAMAGUCHI, N., TAKAHASHI, Y., OHKURA, T., MAEJIMA, Y. & ARAO, T. 2014. Arsenic distribution and speciation near rice roots influenced by iron plaques and redox conditions of the soil matrix. *Environ Sci Technol*.
YANG, H. C., CHENG, J., FINAN, T. M., ROSEN, B. P. & BHATTACHARJEE, H. 2005. Novel pathway for arsenic detoxification in the legume symbiont Sinorhizobium meliloti. *J Bacteriol,* 187**,** 6991-7.
YEDJOU, C., TCHOUNWOU, P., JENKINS, J. & MCMURRAY, R. 2010. Basic mechanisms of arsenic trioxide (ATO)-induced apoptosis in human leukemia (HL-60) cells. *J Hematol Oncol,* 3**,** 28.
YEDJOU, C., THUISSEU, L., TCHOUNWOU, C., GOMES, M., HOWARD, C. & TCHOUNWOU, P. 2009. Ascorbic Acid Potentiation of Arsenic Trioxide Anticancer Activity Against Acute Promyelocytic Leukemia. *Arch Drug Inf,* 2**,** 59-65.
YEDJOU, C. G., MOORE, P. & TCHOUNWOU, P. B. 2006. Dose- and time-dependent response of human leukemia (HL-60) cells to arsenic trioxide treatment. *Int J Environ Res Public Health,* 3**,** 136-40.
YEDJOU, C. G., ROGERS, C., BROWN, E. & TCHOUNWOU, P. B. 2008. Differential effect of ascorbic acid and n-acetyl-L-cysteine on arsenic trioxide-mediated oxidative stress in human leukemia (HL-60) cells. *J Biochem Mol Toxicol,* 22**,** 85-92.
YEDJOU, C. G. & TCHOUNWOU, P. B. 2007. In-vitro cytotoxic and genotoxic effects of arsenic trioxide on human leukemia (HL-60) cells using the MTT and alkaline single cell gel electrophoresis (Comet) assays. *Mol Cell Biochem,* 301**,** 123-30.
YEDJOU, C. G. & TCHOUNWOU, P. B. 2009. Modulation of p53, c-fos, RARE, cyclin A, and cyclin D1 expression in human leukemia (HL-60) cells exposed to arsenic trioxide. *Mol Cell Biochem,* 331**,** 207-14.
YOSHIDA, T., YAMAUCHI, H. & FAN SUN, G. 2004. Chronic health effects in people exposed to arsenic via the drinking water: dose-response relationships in review. *Toxicol Appl Pharmacol,* 198**,** 243-52.
YSART, G., MILLER, P., CREWS, H., ROBB, P., BAXTER, M., DE L'ARGY, C., LOFTHOUSE, S., SARGENT, C. & HARRISON, N. 1999. Dietary exposure estimates of 30 elements from the UK Total Diet Study. *Food Addit Contam,* 16**,** 391-403.
ZAKHARYAN, R. A. & APOSHIAN, H. V. 1999. Enzymatic reduction of arsenic compounds in mammalian systems: the rate-limiting enzyme of rabbit liver

arsenic biotransformation is MMA(V) reductase. *Chem Res Toxicol,* 12**,** 1278-83.
ZARCINAS, B. A., ISHAK, C. F., MCLAUGHLIN, M. J. & COZENS, G. 2004. Heavy metals in soils and crops in Southeast Asia. *Environmental Geochemistry and Health,* 26**,** 343-357.
ZAVALA, Y. J. & DUXBURY, J. M. 2008. Arsenic in rice: I. Estimating normal levels of total arsenic in rice grain. *Environ Sci Technol,* 42**,** 3856-60.
ZAVALA, Y. J., GERADS, R., GORLEYOK, H. & DUXBURY, J. M. 2008. Arsenic in rice: II. Arsenic speciation in USA grain and implications for human health. *Environ Sci Technol,* 42**,** 3861-6.
ZHAO, F. J., HARRIS, E., YAN, J., MA, J., WU, L., LIU, W., MCGRATH, S. P., ZHOU, J. & ZHU, Y. G. 2013a. Arsenic methylation in soils and its relationship with microbial arsM abundance and diversity, and as speciation in rice. *Environ Sci Technol,* 47**,** 7147-54.
ZHAO, F. J., MCGRATH, S. P. & MEHARG, A. A. 2010. Arsenic as a food chain contaminant: mechanisms of plant uptake and metabolism and mitigation strategies. *Annu Rev Plant Biol,* 61**,** 535-59.
ZHAO, F. J., ZHU, Y. G. & MEHARG, A. A. 2013b. Methylated arsenic species in rice: geographical variation, origin, and uptake mechanisms. *Environ Sci Technol,* 47**,** 3957-66.
ZHU, Y. G., WILLIAMS, P. N. & MEHARG, A. A. 2008. Exposure to inorganic arsenic from rice: a global health issue? *Environ Pollut,* 154**,** 169-71.

APPENDICES

Appendix 1: Total Diet Study Foods with Arsenic

TDS Foods With Arsenic

Food Category	Maximum (mg/kg)	TDS Food No	TDS Food Description
baby foods	0.011	227	pears, strained/junior
	0.012	207	chicken, strained/junior, with/without broth or gravy
		232	custard pudding, strained/junior
	0.013	221	sweet potatoes, strained/junior
	0.014	213	vegetables and ham, strained/junior
		224	creamed spinach, strained/junior
	0.015	202	milk-based infant formula, high iron, ready-to-feed
	0.016	313	bananas with tapioca, strained/junior
	0.018	233	fruit dessert/pudding, strained/junior
	0.02	316	split peas with vegetables and ham/bacon, strained/junior
		319	rice cereal with apple, strained/junior
	0.022	212	vegetables and chicken, strained/junior
	0.023	309	soy-based infant formula, ready-to-feed
	0.024	225	applesauce, strained/junior
		230	apple juice, strained
	0.025	214	chicken noodle dinner, strained/junior
		215	macaroni, tomatoes, and beef, strained/junior
	0.026	223	peas, strained/junior
	0.027	222	creamed corn, strained/junior
	0.03	312	rice cereal, strained/junior
	0.034	211	vegetables and beef, strained/junior
	0.04	317	teething biscuits
	0.047	216	turkey and rice, strained/junior
	0.087	311	rice infant cereal, instant, prepared with whole milk
	0.095	220	mixed vegetables, strained/junior
	0.4	218	carrots, strained/junior
beverages	0.003	201	tap water
	0.014	198	beer
	0.038	199	dry table wine
dairy	0.012	8	evaporated milk, canned
	0.02	4	skim milk, fluid
	0.021	237	cream cheese
	0.026	164	butter, regular (salted)
	0.029	236	Swiss cheese
	0.043	235	fruit-flavored yogurt, lowfat (fruit
eggs	0.013	36	eggs, fried
	0.022	37	eggs, boiled
	0.03	35	eggs, scrambled
fats/oils	0.013	301	brown gravy, homemade

Food Category	Maximum (mg/kg)	TDS Food No	TDS Food Description
	0.026	303	Italian salad dressing, low-calorie
	0.086	302	French salad dressing, regular
	0.092	304	olive/safflower oil
	0.115	298	yellow mustard
fruits	0.008	85	pear, raw
	0.009	351	Cranberry juice cocktail,
	0.01	88	grapes, red/green, seedless, raw
	0.011	86	strawberries, raw
		100	grapefruit juice, from frozen
	0.013	307	fruit drink, canned
	0.014	103	prune juice, bottled
		253	apricot, raw
	0.018	93	pineapple, canned in juice
	0.019	94	sweet cherries, raw
		98	orange juice, from frozen concentrate
	0.023	257	grape juice, from frozen concentrate
	0.024	83	peach, raw
		87	fruit cocktail, canned in heavy syrup
	0.025	89	cantaloupe, raw
		256	pineapple juice, from frozen
	0.032	185	apple pie, fresh/frozen, commercial
	0.037	97	avocado, raw
	0.04	96	prunes, dried
		99	apple juice, bottled
		254	peach, canned in light/medium syrup
	0.042	95	raisins, dried
	0.044	78	apple, red, raw
		350	Fruit juice blend (100% juice), canned/bottled
grains	0.011	51	oatmeal, quick (1-3 min), cooked
		60	cornbread, homemade
	0.012	67	corn chips
		68	pancake from mix
		149	spaghetti with tomato sauce, canned
	0.013	252	butter-type crackers
	0.015	347	Spaghetti, enriched, boiled
	0.017	64	rye bread
	0.019	344	Pancakes, frozen, heated
	0.021	59	white roll
		63	tortilla, flour
		73	shredded wheat cereal
		289	chocolate snack cake with chocolate
	0.023	66	saltine crackers
	0.024	184	sandwich cookies with creme filling, commercial
	0.027	291	brownies, commercial
	0.028	369	Cake, yellow w/ icing

Food Category	Maximum (mg/kg)	TDS Food No	TDS Food Description
	0.029	294	pretzels, hard, salted, any shape
	0.031	250	English muffin, plain, toasted
	0.034	62	whole wheat bread
	0.037	74	raisin bran cereal
	0.039	182	sweet roll/Danish, commercial
	0.041	251	graham crackers
	0.043	370	Granola bar, w/ raisins
	0.045	77	oat ring cereal
	0.047	72	fruit-flavored, sweetened cereal
	0.052	248	cracked wheat bread
	0.054	76	granola cereal
		324	BF, cereal, rice, dry, prep w/ water
	0.128	50	white rice, cooked
	0.286	55	corn, canned
	0.32	75	crisped rice cereal
legumes	0.015	42	lima beans, immature, frozen, boiled
	0.02	39	pork and beans, canned
	0.021	38	pinto beans, dry, boiled
	0.022	245	kidney beans, dry, boiled
	0.031	246	peas, mature, dry, boiled
	0.038	247	mixed nuts, no peanuts, dry roasted
	0.081	48	peanuts, dry roasted
	0.086	47	peanut butter, smooth
meat, poultry, fish	0.011	17	ham, baked
		335	Luncheon meat (chicken/turkey)
	0.012	14	beef chuck roast, baked
	0.015	148	meatloaf, homemade
	0.017	13	ground beef, pan-cooked
		239	ham luncheon meat, sliced
	0.018	16	beef steak, loin, pan-cooked
		238	veal cutlet, pan-cooked
	0.021	337	Chicken thigh, oven-roasted (skin removed)
	0.024	328	BF, turkey and broth/gravy
	0.028	152	chicken potpie, frozen, heated
		339	Catfish, pan-cooked w/ oil
	0.029	241	chicken nuggets, fast-food
	0.037	336	Chicken breast, fried, fast-food (w/
	0.04	338	Chicken leg, fried, fast-food (w/ skin)
	0.055	19	pork sausage, pan-cooked
		27	liver, beef, fried
	0.078	26	turkey breast, roasted
	0.081	240	chicken breast, roasted
	0.083	242	chicken, fried (breast, leg, and thigh), fast-food
	0.086	24	chicken, fried (breast, leg, and thigh) homemade
	1.19	318	salmon, steaks or filets, fresh or frozen, baked

Food Category	Maximum (mg/kg)	TDS Food No	TDS Food Description
	1.443	340	Tuna, canned in water, drained
	1.71	32	tuna, canned in oil
	2.68	244	shrimp, boiled
	2.79	34	fish sticks, frozen, heated
	10.4	243	haddock, pan-cooked
mixtures	0.011	274	turkey with gravy, dressing, potatoes, and vegetable, frozen meal, heated
		278	egg, cheese, and ham on English muffin, fast-food
		366	Chicken filet (broiled) sandwich on bun, fast-food
	0.015	331	Meal replacement, liquid RTD, any
		362	Beef w/ vegetables in sauce, from Chinese carry-out
	0.026	275	quarter-pound cheeseburger on bun, fast-food
		363	Chicken w/ vegetables in sauce, from Chinese carry-out
	0.03	271	chili con carne with beans, homemade
	0.032	281	cheese and pepperoni pizza, regular crust, from pizza carry-out
	0.039	282	beef chow mein, from Chinese carry
	0.042	269	beef stroganoff, homemade
	0.044	325	BF, cereal, rice w/apples, dry, prep w/ water
	0.047	284	mushroom soup, canned, condensed, prepared with whole milk
	0.06	270	green peppers stuffed with beef and rice, homemade
	0.063	279	taco/tostada, from Mexican carry-out
	0.09	273	Salisbury steak with gravy, potatoes, and vegetable, frozen meal, heated
	0.106	364	Fried rice, meatless, from Chinese carry-out
	0.244	272	tuna noodle casserole, homemade
	0.279	285	clam chowder, New England, canned, condensed, prepared with whole milk
	1.6	276	fish sandwich on bun, fast-food
sweets	0.018	190	gelatin dessert, any flavor
	0.033	296	jelly, any flavor
	0.107	293	suckers, any flavor
vegetables	0.01	266	turnip, fresh/frozen, boiled
	0.011	110	cabbage, fresh, boiled
		124	summer squash, fresh/frozen, boiled
		358	Sweet potatoes, canned
	0.013	357	Lettuce, leaf, raw
	0.014	108	collards, fresh/frozen, boiled
		109	iceberg lettuce, raw
	0.015	128	onion, mature, raw
		262	beets, fresh/frozen, boiled
	0.016	119	tomato sauce, plain, bottled
		263	Brussels sprouts, fresh/frozen, boiled
		356	Carrot, baby, raw
	0.018	125	green pepper, raw
	0.02	132	radish, raw
	0.022	138	potato chips
		265	eggplant, fresh, boiled
	0.025	123	cucumber, raw

Food Category	Maximum (mg/kg)	TDS Food No	TDS Food Description
		268	mixed vegetables, frozen, boiled
	0.026	140	sweet potato, fresh, baked
		161	dill cucumber pickles
	0.027	258	French fries, fast-food
	0.03	299	black olives
	0.042	137	white potato, baked with skin
	0.043	107	spinach, fresh/frozen, boiled
		259	carrot, fresh, boiled
		267	okra, fresh/frozen, boiled
	0.056	297	sweet cucumber pickles
	0.203	264	mushrooms, raw

Appendix 2: U.S. Total Diet Study Foods without Arsenic

TDS Foods without Arsenic

Food Category	TDS Food No	TDS Food Description
baby foods	203	milk-based infant formula, low iron, ready-to-feed
	205	beef, strained/junior
	208	chicken/turkey with vegetables, high/lean meat, strained/junior
	209	beef with vegetables, high/lean meat, strained/junior
	210	ham with vegetables, high/lean meat, strained/junior
	219	green beans, strained/junior
	226	peaches, strained/junior
	231	orange juice, strained
	310	egg yolk, strained/junior
	314	beets, strained/junior
	320	squash, strained/junior
beverages	191	cola carbonated beverage
	193	fruit drink, from powder
	194	low-calorie cola carbonated beverage
	196	coffee, decaffeinated, from instant
	197	tea, from tea bag
	200	whiskey
	305	coffee, from ground
	306	fruit-flavored carbonated beverage
	308	martini
	380	Bottled drinking water (mineral/spring), not carbonated or flavored
	381	Decaffeinated coffee, from ground
	382	Decaffeinated tea, from tea bag
dairy	1	whole milk, fluid
	2	lowfat (2% fat) milk, fluid
	3	chocolate milk, fluid
	6	plain yogurt, lowfat
	7	chocolate milk shake, fast-food
	10	American, processed cheese
	11	cottage cheese, 4% milkfat
	12	cheddar cheese
	162	margarine, stick, regular (salted)
	166	mayonnaise, regular, bottled
	167	half & half cream
	168	cream substitute, frozen
	175	chocolate pudding, from instant mix
	177	vanilla flavored light ice cream
	286	vanilla ice cream
	300	sour cream
	332	Cottage cheese, creamed, lowfat (2% milk fat)
	333	Sour cream dip, any flavor

Food Category	TDS Food No	TDS Food Description
	368	Pudding, ready-to-eat, flavor other than chocolate
fats/oils	173	tomato catsup
	373	Sweet & sour sauce
	375	Salad dressing, creamy/buttermilk type, regular
	376	Salad dressing, creamy/buttermilk type, low-calorie
	377	Salad dressing, Italian, regular
	378	Olive oil
	379	Vegetable oil
fruits	79	orange, raw
	80	banana, raw
	81	watermelon, raw
	84	applesauce, bottled
	91	plums, raw
	92	grapefruit, raw
	105	lemonade, from frozen concentrate
	186	pumpkin pie, fresh/frozen, commercial
	255	pear, canned in light syrup
	348	Apricots, canned in heavy/light syrup
	352	Orange juice, bottled/carton
grains	52	wheat cereal, farina, quick (1-3min), cooked
	53	corngrits, regular, cooked
	54	corn, fresh/frozen, boiled
	56	cream style corn, canned
	57	popcorn, popped in oil
	58	white bread
	61	biscuit, from refrigerated dough, baked
	65	blueberry muffin, commercial
	69	egg noodles, boiled
	70	macaroni, boiled
	71	corn flakes
	178	chocolate cake with chocolate icing, commercial
	179	yellow cake with white icing, prepared from cake and icing mixes
	183	chocolate chip cookies, commercial
	249	bagel, plain
	290	cake doughnuts with icing, any flavor, from doughnut store
	292	sugar cookies, commercial
	323	BF, cereal, oatmeal, dry, prep w/ water
	345	Breakfast tart/toaster pastry
	346	Macaroni salad, from grocery/deli
	367	Soup, Oriental noodles (ramen noodles), prep w/ water
	372	Popcorn, microwave, butter-flavored
	374	Brown gravy, canned or bottled
legumes	46	green peas, fresh/frozen, boiled
	341	Refried beans, canned
	342	White beans, dry, boiled

Food Category	TDS Food No	TDS Food Description
	343	Sunflower seeds (shelled), roasted,
meat, poultry, fish	18	pork chop, pan-cooked
	20	pork bacon, pan-cooked
	21	pork roast, baked
	22	lamb chop, pan-cooked
	28	frankfurters, beef, boiled
	29	bologna, sliced
	30	salami, sliced
	326	BF, veal and broth/gravy
	327	BF, lamb and broth/gravy
	334	Beef steak, loin/sirloin, broiled
mixtures	142	spaghetti with tomato sauce and meatballs, homemade
	143	beef stew with potatoes, carrots, and onion, homemade
	145	chili con carne, beef and beans, canned
	146	macaroni and cheese, from box mix
	147	quarter-pound hamburger on bun, fast-food
	151	lasagna with meat, homemade
	155	chicken noodle soup, canned, condensed, prepared with water
	157	vegetable beef soup, canned, condensed, prepared with water
	160	white sauce, homemade
	277	frankfurter on bun, fast-food
	280	cheese pizza, regular crust, from pizza carry-out
	283	bean with bacon/pork soup, canned, condensed, prepared with water
	360	Beef and vegetable stew, canned
	361	Lasagna w/ meat, frozen, heated
	365	Burrito w/ beef, beans and cheese, from Mexican carry-out
sweets	169	white sugar, granulated
	170	pancake syrup
	172	honey
	187	milk chocolate candy bar, plain
	188	caramel candy
	287	fruit flavor sherbet
	288	popsicle, any flavor
	295	chocolate syrup dessert topping
	371	Candy bar, chocolate, nougat, and nuts
vegetables	111	coleslaw with dressing, homemade
	112	sauerkraut, canned
	113	broccoli, fresh/frozen, boiled
	114	celery, raw
	115	asparagus, fresh/frozen, boiled
	116	cauliflower, fresh/frozen, boiled
	117	tomato, red, raw
	121	green beans, fresh/frozen, boiled
	122	beans, snap green, canned
	126	winter squash, fresh/frozen, baked, mashed

Food Category	TDS Food No	TDS Food Description
	131	beets, canned
	134	French fries, frozen, heated
	135	mashed potatoes, from flakes
	136	white potato, boiled without skin
	139	scalloped potatoes, homemade
	156	tomato soup, canned, condensed, prepared with water
	260	tomato, stewed, canned
	261	tomato juice, bottled
	353	Potato salad, mayonnaise-type, from grocery/deli
	354	Potato, mashed, prepared from fresh
	355	Coleslaw, mayonnaise-type, from grocery/deli
	359	Tomato salsa, bottled

Appendix 3: Concentration of arsenic in rice samples

Product Category	Inorganic As per serving (mcg/serving)	Country of Origin	Sample Description	Sample ID	
Basmati Rice	3	USA	Organic Brown Basmati	242778	1
	3.1	India	Aged	728649	1
	3.2	India	Imported Aged	721863A	1
	3.3	India	White Basmati Rice	728645	1
	3.5	India	White Basmati Rice	728653	1
	3.7	India	Basmati Rice	720020	1
			Quick-Cook Organic Brown, parboiled	728646	1
	3.9	India	Organic Aged	721867	1
			Organic White Basmati Rice	721857B	1
			White Basmati Rice	721858B	1
		USA	California Brown	720013	1
	4	India	Organic	728644	1
	4.1	USA	California Brown	720130	1
	4.7	Not Determined	Organic White Basmati	721866	1
		USA	Organic Brown Basmati	720112	1
	4.8	USA	Basmati Rice	720019	1
	5.3	India	Brown Basmati Rice	728654	1
	5.6	India	Brown Basmati Rice	728647	1
	5.7	India	Brown Basmati Rice	721855B	1
	6	India	Brown	720016	1
	6.5	India	Basmati Rice	720018	1
	6.6	USA	Organic Long Grain Brown Basmati	720014	1
	9	USA	Basmati Brown Rice	720110	1
			Brown	721867B	1
Rice (non-Basmati)	3	USA	Instant Enriched Long Grain	492961	1
	3.2	Not Determined	Long Grain White Rice Fully Cooked, parboiled	728638	1
	3.3	Not Determined	Ready to Serve Long Grain White Rice Fully Cooked, parboiled	492963	1
	3.6	Not Determined	Whole Grain Brown Pre-cooked parboiled	725130	1
	4	USA	Whole Grain Red	721856A	1
	4.2	Italy	Organic Carnaroli	721866B	1
	4.4	USA	Medium Grain Rice	720004	1
	4.7	Not Determined	Enriched Rice Long Grain	725139	1
	5.1	USA	Long Grain Brown	721857	1
	5.4	Not Determined	Enriched Long Grain	720000	1
		USA	Extra Long Grain	720003	1
			Long Grain Brown	721860B	1
	5.8	USA	Long Grain White	70139	1
	5.9	USA	Enriched Long Grain Rice	720005	1
			Extra Long Grain Rice	719998	1
			Wild Rice	721856B	1
	6	USA	Wild Riced	721868A	1
	6.2	USA	Cracked Wild Rice	721868B	1
			Long Grain Rice	242774	1
				719997	1
			Medium Grain	720007	1
	6.5	USA	Enriched Extra Long Grain	492959	1
	6.6	USA	Extra Long Grain	70137	1
			Whole Grain Brown Rice	720122	1
	6.8	Not Determined	Long Grain Rice	242777	1
		USA	Extra Long Grain	720008	1
				721877	1
	6.9	USA	Extra Long Grain Rice	721869	1
	7.1	USA	Xtra Long Grain Rice	720124	1
	7.4	USA	Enriched Parboiled Long Grain	492960	1
			Long Grain	70138	1
	7.6	Not Determined	Brown Rice Natural Long Grain	492962	1
	7.9	USA	Whole Grain Brown Rice	720123	1
	8	USA	Extra Long Grain	720010	1
			Whole Grain Brown	70134	1
	8.2	USA	Brown Rice Whole Grain	720125	1
	8.4	USA	Brown Natural Whole Grain	719999	1
	8.8	USA	Whole Grain, Long Grain	492958	1
	9.1	USA	Long Grain Brown Whole	720009	1
	9.4	USA	Long Grain Brown Rice Extra Fancy	721864A	1
			Whole Grain Brown Rice	720120	1
	9.7	Not Determined	Brown Rice Natural Long Grain	725137	1
	10.2	Not Determined	Long Grain Brown Rice	720006	1
	10.4	USA	Long Grain Brown Rice	70136	1
			Whole Grain Brown Rice	720121	1
	10.5	USA	Long Grain Natural Brown Rice	720001	1
	11.1	USA	Brown Natural Whole Grain	70135	1

Product Category	Inorganic As per serving (mcg/serving)	Country of Origin	Sample Description	Sample ID	
Rice Beverage	3.4	Not Reported	Rice Drink Organic Original	721854A	1
	3.8	Not Reported	Organic	719980	1
	4.1	Not Reported	Organic	719976	1
Rice Cakes	3	Not Reported	Organic Unsalted	719993	1
				719994	1
			Wild Organic Gluten-free Vegan Lightly Salted	721871	1
	3.1	Not Reported	Brown Lightly Salted	728639	1
			Caramel Corn Gluten-free	720119	1
			Organic Brown	721878	1
	3.2	Not Reported	Organic Brown	728640	1
	3.7	Not Reported	Brown Rice Organic Salt-free	719991	1
	3.8	Not Reported	Brown Rice Organic Salt-free	721858	1
	4	Not Reported	Butter Popped Corn	720116	1
	4.6	Not Reported	Lightly Salted	721876	1
	4.7	Not Reported	Brown Rice Whole Grain Lightly Salted	242776	1
	4.8	Not Reported	Brown Salt Free	721862A	1
	5	Not Reported	Apple Cinnamon	720118	1
	5.1	Not Reported	Apple Cinnamon	719997B	1
	5.6	Not Reported	Butter Popped Corn	720117B	1
			Salt Free	721865B	1
	6.1	Not Reported	Salt Free	728642	1
	6.2	Not Reported	Lightly Salted	719996	1
			Salt Free	70133	1
	6.3	Not Reported	Salt Free	70131	1
				70132	1
	6.5	Not Reported	Plain Salt Free Whole Grain Brown	728641	1
	6.6	Not Reported	Plain Salt Free Whole Grain Brown	725135	1
	6.8	Not Reported	Salt Free, whole grain brown rice	721865	1
	7.2	Not Reported	Low Sodium	719995	1
	7.6	Not Reported	Salted Plain Gluten Free	728643	1
	7.7	Not Reported	Lightly Salted	719992	1
	8	Not Reported	Salt Free	242773	1
				720115B	1
	8.2	Not Reported	Salt Free	720115A	1
			Sodium Free Plain Gluten Free	725136	1
Rice Cereal	3.2	Not Reported	Crispy Brown Rice Organic	492957	1
			Crispy Rice Gluten Free Whole Grain Brown	70146	1
			Crispy Rice Gluten-free Whole Grain Brown Rice	725131	1
			Crispy Rice Toasted	70142	1
			Organic Whole Grain Rice Cereal (infant)	70145	1
	3.4	Not Reported	Crispy Brown Rice Gluten Free Organic	721859	1
	3.6	Not Reported	Crispy Rice Toasted	725134	1
	3.8	Not Reported	Crispy Rice	719989	1
			Rice Cereal	242770	1
	3.9	Not Reported	Crispy Rice	719984	1
			Crispy Rice Toasted	70143	1
				70149	1
			Crispy Toasted Rice	719985	1
	4.1	Not Reported	Crisp Rice Toasted	721869B	1
			Honey Rice Gluten Free	719982	1
	4.2	Not Reported	Crispy Rice Toasted	721874	1
	4.3	Not Reported	Rice Chex Gluten-free	725132	1
	7.1	Not Reported	Organic Brown Rice Crisps	70144	1
	7.3	Not Reported	Organic Brown Rice Crisps	720342	1
	9.7	Not Reported	Gluten Free Cream of Brown Rice	719981	1

Appendix 4: Informative Sentences on Arsenic Concentration in Rice

Sentence Identifier	Sentence Text	PMID
908317_9	Much lower levels were found in all the other food types analyzed; of these, the highest levels found were a mean level of 0.08 ppm in chicken and 0.16 ppm in rice.	908317
7806208_6	The method was also applied to the determination of arsenic in reference material rice powder NBS 1568 and the result, 0.37-0.43mg/kg, coincided with the given value 0.41 +/- 0.05mg/kg.	7806208
10506007_11	The highest inorganic arsenic values were found in raw rice (74 ng/g), followed by flour (11 ng/g), grape juice (9 ng/g) and cooked spinach (6 ng/g).	10506007
11891266_6	High-affinity uptake (0-0.0532 mM) for arsenite and arsenate with eight rice varieties, covering two growing seasons, rice var.	11891266
11918027_8	Concentrations of arsenic in rice grain did not exceed the food hygiene concentration limit (1.0 mg of As kg(-1) dry weight).	11918027
11918027_9	The concentrations of arsenic in rice straw (up to 91.8 mg kg(-1) for the highest As treatment) were of the same order of magnitude as root arsenic concentrations (up to 107.5 mg kg(-1)), suggesting that arsenic can be readily translocated to the shoot.	11918027
12059152_6	Their levels were, in general, below the maximum limits establish by the Spaniard legislation; however, the As concentration in the licorice sticks was above this maximum limit (0.11 +/- 0.01 microg g(-)(1)).	12059152
12059152_9	The As concentration in the licorice extract was 0.503 +/- 0.01 microg g(-)(1), and could represent a serious hazard to human health if it is used in high proportions.	12059152
12564892_10	Rice grain grown in the regions where arsenic is building up in the soil had high arsenic concentrations, with three rice grain samples having levels above 1.7 microg g(-1).	12564892
12635819_5	Mean values for the boro and aman season rices were 183 and 117 microg/kg, respectively.	12635819
12635819_9	Human exposure to arsenic through rice would be equivalent to half of that in water containing 50 microg/kg for 14% of the paddy rice samples at rice and water intake levels of 400 g and 4 L/cap/day, respectively.	12635819
12734624_4	The detection limits for dry flour rice expressed as As were 2 and 3 ng g(-1) for As(III) and AsB on the cation column and 3, 6 and 5 ng g(-1) for As(V), MMA and DMA, respectively, on the anion column.	12734624
12806106_8	As is highly elevated in rice leaves from the Dukpyung (1.14 mg kg(-1)) and the Chubu areas (1.35 mg kg(-1)).	12806106
14987870_7	However, rice plants, especially the roots had a significantly higher concentration of arsenic (2.4 mg/kg) compared to stem (0.73 mg/kg) and rice grains (0.14 mg/kg).	14987870
15084107_5	Arsenic concentrations found in rice-based cereals (63-320 ng/g dry weight) were similar to those reported for raw rice.	15084107
15084107_7	Arsenic content in puree infant food products, including rice cereals, fruits, and vegetables, varies from <1 to 24 ng/g wet weight.	15084107
15234998_3	After iron plaque on rice roots was induced in solutions containing 20, 40, 60, 80, and 100 mg Fe2+ l(-1), seedlings were transplanted into nutrient solution with 0.5 mg As l(-1).	15234998
15979720_6	High As concentrations were found in the rice grain that ranged from 0.5 to 7.5 mg/kg, most of which exceed the maximal permissible limit of 1.0 mg/kg dry matter.	15979720
16003581_4	The concentration of the total arsenic in the samples i.e. water (n = 64), soil (n = 30), sediment (n = 27) and rice grain (n = 10) were ranged from 15 to 825 microg L(-1), 9 to 390 mg kg(-1), 19 to 489 mg kg(-1) and 0.018 to 0.446 mg kg(-1), respectively	16003581
16082952_9	The dimethylarsinic acid (DMA) and inorganic arsenic concentration ranged from 22 to 270 ng of As/g of rice and from 31 to 108 ng of As/g of rice, respectively, for samples cooked in reagent water.	16082952
16124284_6	USA long grain rice had the highest mean arsenic level in the grain at 0.26 microg As g(-1) (n = 7), and the highest grain arsenic value of the survey at 0.40 microg As g(-1).	16124284
16124284_7	The mean arsenic level of Bangladeshi rice was 0.13 microg As g(-1) (n = 15).	16124284
16730050_6	Digestion and analysis of individual grains of boro winter rice from the 2 sites irrigated with groundwater containing 150 and 180 microg/L As yielded concentrations of 0.28+/-0.13 mg/kg (n=12) and 0.44+/-0.25 mg/kg (n=12), respectively.	16730050
16730050_7	The As content of winter rice from the control site was not significantly different though less variable (0.30+/-0.07; n=12).	16730050
16839594_7	Arsenic concentrations were 0.40+/-0.03 and 0.58+/-0.12 mg/kg in parboiled rice of arsenic affected area, cooked with excess water and 1.35+/-0.04 and 1.59+/-0.07 mg/kg in gruel for BRRI dhan28 and BRRI hybrid dhan1, respectively.	16839594
16839594_8	In non-parboiled rice, arsenic concentrations were 0.39+/-0.04 and 0.44+/-0.03 mg/kg in rice cooked with excess water and 1.62+/-0.07 and 1.74+/-0.05 mg/kg in gruel for BRRI dhan28 and BRRI hybrid dhan1, respectively.	16839594
16839594_9	Total arsenic content in rice, cooked with limited water (therefore gruel was absorbed completely by rice) were 0.89+/-0.07 and 1.08+/-0.06 mg/kg (parboiled) and 0.75+/-0.04 and 1.09+/-0.06 mg/kg (non-parboiled) for BRRI dhan28 and BRRI hybrid dhan1, resp	16839594
16839594_10	Water used for cooking rice contained 0.13 and 0.01 mg of As/l for contaminated and non-contaminated areas, respectively.	16839594
16875714_7	The mean total arsenic level in 46 rice samples was 358 microg/kg (range: 46 to 1,110 microg/kg dry weight) and 333 microg/kg (range: 19 to 2,334 microg/kg dry weight) in 39 vegetable samples.	16875714

Sentence Identifier	Sentence Text	PMID
16875714_10	Using individual, self-reported data on daily consumption of rice and drinking water the total arsenic ADI was 1,176 microg (range: 419 to 2,053 microg), 14% attributable to inorganic arsenic in cooked rice.	16875714
16876928_5	Using low-arsenic water (As < 3 microg/L), the traditional method of the Indian subcontinent (wash until clear; cook with rice: water::1:6; discard excess water) removed up to 57% of the arsenic from rice containing arsenic 203-540 microg/kg.	16876928
16955884_5	The districts with the highest mean arsenic rice grain levels were all from southwestern Bangladesh: Faridpur (boro) 0.51 > Satkhira (boro) 0.38 > Satkhira (aman) 0.36 > Chuadanga (boro) 0.32 > Meherpur (boro) 0.29 microg As g(-1).	16955884
16955884_9	Daily consumption of rice with a total arsenic level of 0.08 microg As g(-1) would be equivalent to a drinking water arsenic level of 10 microg L(-1).	16955884
17067657_3	In the unaffected areas, where irrigation water contained little As (<1 microg/L), As concentrations of rice field soils ranged from 1.5 to 3.0 mg/kg and did not vary significantly with either depth or sampling time throughout the irrigation period.	17067657
17239924_6	No rice plant survived up to maturity stage in soil treated with 60 and 90 mg of As kg(-1).	17239924
17239924_8	The content of photosynthetic pigments in these five rice varieties did not differ significantly (p>0.05) from each other in control treatment though they differed significantly (p<0.05) from each other in 30 mg of As kg(-1) soil treatment.	17239924
17311403_4	The bioreporter cells detected arsenic in all rice varieties tested, with averages of 0.02-0.15 microg of arsenite equivalent per gram of dry weight and a method detection limit of 6 ng of arsenite per gram of dry rice.	17311403
17346792_6	Arsenic concentration in rice grain was 0.5+/-0.02 mg kg(-1) with the highest concentrations being in grains grown on soil treated with 40 mg As kg(-1) soil.	17346792
17346792_7	With the average rice consumption between 400 and 650 g/day by typical adults in the arsenic-affected areas of Bangladesh, the intake of arsenic through rice stood at 0.20-0.35 mg/day.	17346792
17346792_9	Moreover, when the rice plant was grown in 60 mg of As kg(-1) soil, arsenic concentrations in rice straw were 20.6+/-0.52 at panicle initiation stage and 23.7+/-0.44 at maturity stage, whereas it was 1.6+/-0.20 mg kg(-1) in husk.	17346792
17366772_7	Irrigating a rice field with groundwater containing 0.55 mg/L of arsenic with a water requirement of 1,000 mm results in an estimated addition of 5.5 kg of arsenic per ha per annum.	17366772
17438760_11	Modeling arsenic intake for the U.S. population based on this survey shows that for certain groups (namely Hispanics, Asians, sufferers of Celiac disease, and infants) dietary exposure to inorganic As from elevated levels in rice potentially exceeds the m	17438760
17599387_8	Arsenic concentrations in parboiled and non-parboiled brown rice of BRRI dhan28 were 0.8+/-0.1 and 0.5+/-0.0 mg kg(-1) dry weight, respectively while those of BRRI hybrid dhan1 were 0.8+/-0.2 and 0.6+/-0.2 mg kg(-1) dry weight, respectively.	17599387
17599387_9	However, parboiled and non-parboiled polish rice grain of BRRI dhan28 contained 0.4+/-0.0 and 0.3+/-0.1 mg kg(-1) dry weight of arsenic, respectively while those of BRRI hybrid dhan1 contained 0.43+/-0.01 and 0.5+/-0.0 mg kg(-1) dry weight, respectively.	17599387
17599387_11	The concentration of arsenic found in the present study is much lower than the permissible limit in rice (1.0 mg kg(-1)) according to WHO recommendation.	17599387
17599387_12	Thus, rice grown in soils of Bangladesh contaminated with arsenic of 14.5+/-0.1 mg kg(-1) could be considered safe for human consumption.	17599387
17852383_6	The results show that the arsenic concentration in cooked rice is always higher than that in raw rice and range from 227 to 1642 microg kg(-1).	17852383
17852383_9	The daily inorganic arsenic intakes for water plus rice were 229, 1024 and 2000 microg day(-1) for initial arsenic concentrations in the cooking water of 50, 250 and 500 microg arsenic l(-1), respectively, compared with the tolerable daily intake which is	17852383
17969706_6	Rice grain As levels over 0.60 microg g(-1) d. wt were found in rice grown in paddy soil of around only 10 microg g(-1) As, showing that As in paddy soils is problematic with respect to grain As levels.	17969706
17969706_9	In rice, the export of As from the shoot to the grain appears to be under tight physiological control as the grain/shoot ratio decreases by more than an order of magnitude (from approximately 0.3 to 0.003 mg/kg) and as As levels in the shoots increase fro	17969706
18068879_5	The mean arsenic content of edible plant material (dry weight) were found in the order of onion leaves (0.55 mg As kg(-1))>onion bulb (0.45 mg As kg(-1))>cauliflower (0.33 mg As kg(-1))>rice (0.18 mg As kg(-1))>brinjal (0.09 mg As kg(-1))>potato (<0.01 mg	18068879
18339463_3	Analysis of UK baby rice revealed a median inorganic arsenic content (n=17) of 0.11 mg/kg.	18339463
18339463_6	It was found that 35% of the baby rice samples analysed would be illegal for sale in China which has regulatory limit of 0.15 mg/kg inorganic arsenic.	18339463
18385862_3	All rice milk samples analysed in a supermarket survey (n = 19) would fail the EU limit with up to 3 times this concentration recorded, while out of the subset that had arsenic species determined (n = 15), 80% had inorganic arsenic levels above 10 microg	18385862
18503245_9	When the method was applied to ten short-grain brown rice samples, the iAs concentrations were 0.108-0.227 mg/kg dry weight and the total As concentrations were 0.118-0.260 mg/kg dry weight.	18503245
18546734_5	Total As concentration in rice varied from 0.005 to 0.710 mg kg(-1).	18546734
18546734_6	We combined our data set with literature values to derive a global "normal" range of 0.08-0.20 mg kg(-1) for As concentration in rice.	18546734
18546734_7	The mean As concentrations for rice from the U.S. and Europe (both 0.198 mg kg(-1)) were statistically similar and significantly higher than rice from Asia (0.07 mg kg(-1)).	18546734

Sentence Identifier	Sentence Text	PMID
18546734_9	Wide variability found in U.S. rice grain was primarily influenced by region of growth rather than commercial type, with rice grown in Texas and Arkansas having significantly higher mean As concentrations than that from California (0.258 and 0.190 versus	18546734
18546735_8	DMA increased linearly with increasing total As but arsenite remained fairly constant at approximately 0.1 mg kg(-1), showing that rice high in As was dominated by DMA.	18546735
18546736_4	The analyses were performed in raw rice and in rice cooked by boiling to dryness in water spiked with As(V) (0.1-1 microg mL(-1)).	18546736
18546736_5	In raw rice, inorganic As represented 27-93% of total As: total As = 0.188 +/- 0.078 microg g(-1) dry weight (dw); inorganic As = 0.114 +/- 0.046 microg g(-1) dw.	18546736
18546736_6	After cooking, the rice retained between 45% and 107% of the As(V) added to the cooking water, and the inorganic As concentrations ranged between 0.428 microg g(-1) dw (0.1 microg mL(-1) in the cooking water) and 3.89 microg g(-1) dw (1.0 microg microL(-1	18546736
18546736_7	For raw rice, the inorganic As intake of the Spanish population (16 g raw rice/day) remains below the tolerable daily intake (TDI) proposed by the WHO (2.1 microg inorganic As/day/kg body weight).	18546736
18602205_13	The daily dietary intake of inorganic arsenic (mug/kg body wt./day) by an adult from rice grain itself (2.32) is higher than the WHO recommended PTDI value of inorganic arsenic (2.1) and inorganic arsenic contributes 96.8% of the total dietary intakes of	18602205
18637329_11	Arsenic concentrations in rice grain are lower than the food safety limitation in China (0.7 mg x kg(-1)).	18637329
18644665_8	Soil As (both total and extractable As) was significantly and positively correlated with rice grain As (0.296+/-0.063 microg g(-1), n=56).	18644665
18666619_9	The average contamination of rice with arsenic is 0.57-0.69 mg/kg, which means the intake of this element on the level 30%-45% PTWI (Provisional Tolerable Weekly Intake).	18666619
18678041_7	The vast majority (85%) of the market rice grains possessed total As levels < 150 ng g(-1).	18678041
18678041_8	The rice collected from mine-impacted regions, however, were found to be highly enriched in As, reaching concentrations of up to 624 ng g(-1).	18678041
18678041_11	The mean baseline concentrations for As(i) in Chinese market rice based on this survey were estimated to be 96 ng g(-1) while levels in mine-impacted areas were higher with ca. 50% of the rice in one region predicted to fail the national standard.	18678041
18763528_7	The average arsenic contents in brown rice and husks were 165.1 microg/kg and 144.2 microg/kg, which was also lower than the Chinese Foods Quality Standard.	18763528
18939599_5	Five rice bran solubles products were tested, sourced from the United States and Japan, and were found to have 0.61-1.9 mg/kg inorganic arsenic.	18939599
18939599_6	Manufactures recommend approximately 20 g servings of the rice bran solubles per day, which equates to a 0.012-0.038 mg intake of inorganic arsenic.	18939599
18939599_9	At the manufacturers recommended rice bran solubles consumption rate, inorganic arsenic intake exceeds 0.01 mg/ day, remembering that rice bran solubles are targeted at malnourished children and that actual risk is based on mg kg(-1) day(-1) intake.	18939599
19004533_5	Also, estimated inorganic As intake from groundwater and rice were over Provisional Tolerable Weekly Intake (15 microg/week/kg body wt.) by FAO/WHO for 92% of the residents examined.	19004533
19142738_6	The median and mean total As contents in 214 rice grain samples were 131 and 143 microg/kg, respectively, with a range of 2-557 microg/kg (dry weight, dw).	19142738
19142738_8	Daily dietary intake of As from rice was 56.4 microg for adults (males and females) while the total daily intake of As from rice and from drinking water was 888.4 and 706.4 microg for adult males and adult females, respectively.	19142738
19350943_4	Median total arsenic contents of rice varied 7-fold, with Egypt (0.04 mg/kg) and India (0.07 mg/kg) having the lowest arsenic content while the U.S. (0.25 mg/kg) and France (0.28 mg/kg) had the highest content.	19350943
19534157_6	In contaminated area, daily intake only from cooked Boro rice for 34.6% of the samples exceeded the WHO recommended MTDI value (2 microg In-As day(-1) kg(-1) body wt), whereas daily intake from Aman rice was below MTDI value as was rice from uncontaminate	19534157
19534157_7	Our study indicated that employing traditional rice cooking method as followed in Bengal delta and using water having arsenic <3 microg L(-1) for cooking, actual exposure to arsenic from rice would be much less.	19534157
19634447_5	Rice protreated with -P showed As toxicity symptoms after being exposed to 50 micromol/L arsenate for 4 h, while +P rice did not show any toxicity symptoms.	19634447
19680842_4	Cooking water and raw atab and boiled rice contained 40 microg As l(-1) and 185 and 315 microg As kg(-1), respectively.	19680842
19680842_6	Based on the best possible scenario (the traditional cooking method leading to the lowest level of contamination, and the atab rice type with the lowest As content), t-As daily intake was estimated to be 328 microg, which was twice the tolerable daily int	19680842
19948354_3	Two-week-old rice seedlings were exposed to two concentrations of arsenate (50 or 100 microM), and leaf samples were collected 4d after treatment.	19948354
20071009_6	In the next five days, these volunteers switched to a rice diet, increasing the average arsenic daily intake to 36.4+/-2.8microg and 34.1+/-7.7microg, respectively.	20071009

Appendix 5: Dataset of arsenic content of rice samples constructed from literature curation.

PMID	Sample	Location	Arsenic Content	Year of Publication	Arsenic Species
11918027	Rice straw	Bangladesh	up to 91.8 mg kg(-1)	2002	Arsenic
12059152	Licorice sticks	Spain	0.503+/-0.01 microg (g)(1)	2002	Arsenic
12564892	Rice grains	Bangladesh	>1.7 micro g(-1)	2003	Arsenic
14987870	Rice roots	Bangladesh	2.4 mg/kg	2004	Arsenic
14987870	Rice stem	Bangladesh	0.73 mg/kg	2004	Arsenic
14987870	Rice grains	Bangladesh	0.14 mg/kg	2004	Arsenic
15084107	Rice-based cereals	USA	63-320 ng/g dry weight	2004	Arsenic
15979720	Rice grain	China - Industrial District in Chenzhou, Southern China	0.5 to 7.5 mg/kg	2005	Arsenic
16003581	Rice grain	India - Central India (Ambagarh, Chauki, Chhattisgarh)	0.018 to 0.446 mg kg(-1)	2005	Total Arsenic
16082952	Rice	USA	22 to 270 ng	2005	Dimethylarsinic acid
16082952	Rice	USA	31 to 108 ng	2005	Inorganic arsenic
16730050	Boro Winter Rice irrigated with groundwater containing 150 microg/L arsenic	Bangladesh	0.28+/-0.13 mg/kg	2006	Arsenic
16730050	Boro Winter Rice irrigated with groundwater containing 180 microg/L arsenic	Bangladesh	0.44+/-0.25 mg/kg	2006	Arsenic
16839594	Parboiled Rice of Affected Area for BRRI dhan28, cooked with excess water	Bangladesh	0.40+/-0.03 mg/kg	2006	Arsenic
16839594	Parboiled Rice of Affected Area for BRRI hybrid dhan1, cooked with excess water	Bangladesh	0.58+/-0.12 mg/kg	2006	Arsenic
16839594	Parboiled Rice Gruel for BRRI dhan28	Bangladesh	1.35+/-0.04 mg/kg	2006	Arsenic
16839594	Parboiled Rice Gruel for BRRI hybrid dhan1	Bangladesh	1.59+/-0.07 mg/kg	2006	Arsenic
16839594	Non-Parboiled Rice of Affected Area for BRRI dhan28, cooked with excess water	Bangladesh	0.39+/-0.04 mg/kg	2006	Arsenic

PMID	Sample	Location	Arsenic Content	Year of Publication	Arsenic Species
16839594	Non-Parboiled Rice of Affected Area for BRRI hybrid dhan1, cooked with excess water	Bangladesh	0.44+/-0.03 mg/kg	2006	Arsenic
16839594	Non-Parboiled Rice Gruel for BRRI dhan28	Bangladesh	1.62+/-0.07 mg/kg	2006	Arsenic
16839594	Non-Parboiled Rice Gruel for BRRI hybrid dhan1	Bangladesh	1.74+/-0.05 mg/kg	2006	Arsenic
17346792	Rice grains grown on soil treated with 40 mg As kg(-1)	Bangladesh	0.5+/-0.02 mg kg(-1)	2008	Arsenic
17346792	Straw of rice plant at panicle initiation stage grown on soil treated with 60 mg As kg(-1)	Bangladesh	20.6+/-0.52 mg kg(-1)	2008	Arsenic
17346792	Straw of rice plant at maturity stage grown on soil treated with 60 mg As kg(-1)	Bangladesh	23.7+/-0.44 mg kg(-1)	2008	Arsenic
17346792	Hust of rice plant grown on soil treated with 60 mg As kg(-1)	Bangladesh	1.6+/-0.20 mg kg(-1)	2008	Arsenic
17599387	Parboiled brown rice of BRRI dhan28	Bangladesh	0.8+/-0.1 mg kg(-1) dry weight	2007	Arsenic
17599387	Non-Parboiled brown rice of BRRI dhan28	Bangladesh	0.5+/-0.0 mg kg(-1) dry weight	2007	Arsenic
17599387	Parboiled brown rice of BRRI dhan1	Bangladesh	0.8+/-0.2 mg kg(-1) dry weight	2007	Arsenic
17599387	Non-Parboiled brown rice of BRRI dhan1	Bangladesh	0.6+/-0.2 mg kg(-1) dry weight	2007	Arsenic
17852383	Cooked rice	India - Rural Village of West Bangal, India	227 to 1642 microg kg(-1)	2008	Arsenic
17852383	Daily inorganic arsenic intake for water plus rice for initial arsenic concentrations in cooking water of 50 microg arsenic l(-1)	India - Rural Village of West Bangal, India	229 microg kg(-1)	2008	Arsenic
17852383	Daily inorganic arsenic intake for water plus rice for initial arsenic concentrations in cooking water of 250 microg arsenic l(-1)	India - Rural Village of West Bangal, India	1024 microg kg(-1)	2008	Arsenic
17852383	Daily inorganic arsenic intake for water plus rice for initial arsenic concentrations in cooking water of 500 microg arsenic l(-1)	India - Rural Village of West Bangal, India	2000 microg kg(-1)	2008	Arsenic
18385862	Rice milk samples	United Kingdom	>10 microg l(-1)	2008	Inorganic arsenic

PMID	Sample	Location	Arsenic Content	Year of Publication	Arsenic Species
18503245	Ten short-grain brown rice samples	Japan	0.108-0.227 mg/kg dry weight	2008	Inorganic arsenic
18503245	Ten short-grain brown rice samples	Japan	0.118-0.260 mg/kg dry weight	2008	Total arsenic
18546734	Rice	New York (upstate) supplemented with samples from Canada, France, Venezuela and other countries	0.005 to 0.710 mg kg(-1)	2008	Total arsenic
18546734	Rice Global “normal” range	New York supplemented with samples from Canada, France, Venezuela and other countries	0.08 to 0.20 mg kg(-1)	2008	Total arsenic
18546734	Rice	USA and Europe	0.198 mg kg(-1)	2008	Total arsenic
18546734	Rice	Asia	0.07 mg kg(-1)	2008	Total arsenic
18546734	Rice	USA - Texas, USA	0.258 mg kg(-1)	2008	Total arsenic
18546734	Rice	USA - Arkansas, USA	0.190 mg kg(-1)	2008	Total arsenic
18546734	Rice	USA - California, USA	0.133 mg kg(-1)	2008	Total arsenic
18546736	Rice in 0.1 microg mL(-1) of As(V) in the cooking water	Spain	0.428 microg g(-1) dw	2008	Inorganic arsenic
18546736	Rice in 1.0 microg mL(-1) of As(V) in the cooking water	Spain	3.89 microg g(-1) dw	2008	Inorganic arsenic
18637329	Rice grain	China - Hebei, China	<0.7 mg x kg(-1)	2008	Total arsenic
18678041	Rice collected from mine-impacted regions	China	up to 624 ng g(-1)	2008	Inorganic arsenic
18678041	Rice mean baseline inorganic arsenic concentration in Chinese market rice	China	96 ng g(-1)	2008	Inorganic arsenic

Appendix 6: Articles citing a manuscript from this thesis research

An article from this thesis is:

Johnson, M. O., Cohly, H. H., Isokpehi, R. D., & Awofolu, O. R. (2010). The case for visual analytics of arsenic concentrations in foods. International journal of environmental research and public health, 7(5), 1970-1983.

According to Google Scholar, as of January 2014, the article has been cited 9 times.

Sunita, M. S. L., Prashant, S., Chari, P. B., Rao, S. N., Balaravi, P., & Kishor, P. K. (2012). Molecular identification of arsenic-resistant estuarine bacteria and characterization of their ars genotype. Ecotoxicology, 21(1), 202-212.

Sims, J. N., Isokpehi, R. D., Cooper, G. A., Bass, M. P., Brown, S. D., St John, A. L., ... & Cohly, H. H. (2011). Visual analytics of surveillance data on foodborne vibriosis, United States, 1973–2010. Environmental health insights, 5, 71.

Simmons, S. S., Isokpehi, R. D., Brown, S. D., McAllister, D. L., Hall, C. C., McDuffy, W. M., ... & Cohly, H. H. (2011). Functional Annotation Analytics of *Rhodopseudomonas palustris* Genomes. Bioinformatics and biology insights, 5, 115.

Udensi, U. K., Graham-Evans, B. E., Rogers, C., & Isokpehi, R. D. (2011). Cytotoxicity patterns of arsenic trioxide exposure on HaCaT keratinocytes. Clinical, cosmetic and investigational dermatology, 4, 183.

Isokpehi, R. D., Udensi, U. K., Anyanwu, M. N., Mbah, A. N., Johnson, M. O., Edusei, K., ... & Awofolu, O. R. (2012). Knowledge Building Insights on Biomarkers of Arsenic Toxicity to Keratinocytes and Melanocytes. Biomarker insights, 7, 127.

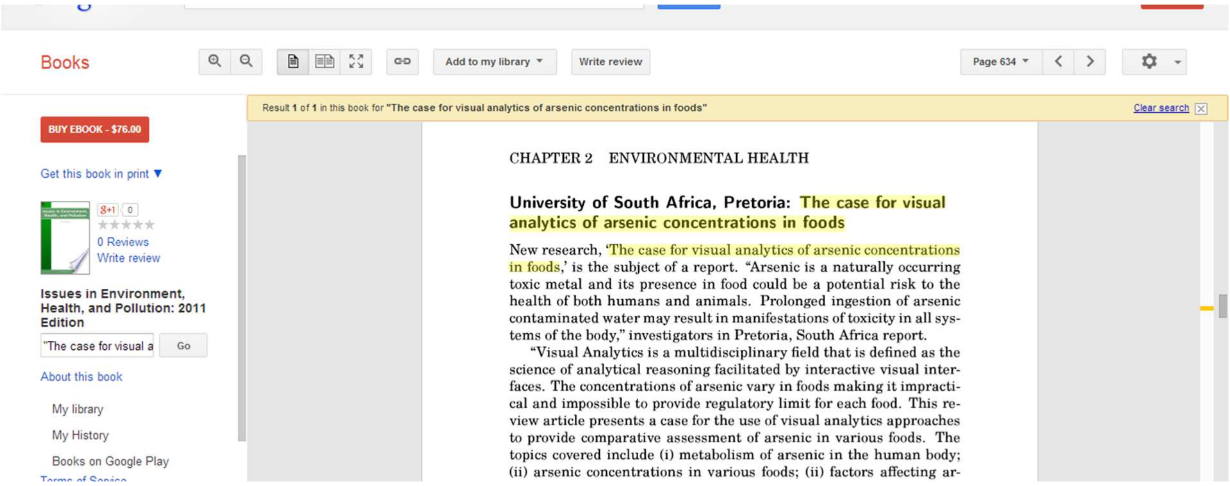
Besante, J., Niforatos, J., & Mousavi, A. (2011). Cadmium in Rice: Disease and Social Considerations. Environmental Forensics, 12(2), 121-123.

Williams, B. S., Isokpehi, R. D., Mbah, A. N., Hollman, A. L., Bernard, C. O., Simmons, S. S., ... & Garner, B. L. (2012). Functional Annotation Analytics of *Bacillus* Genomes Reveals Stress Responsive Acetate Utilization and Sulfate Uptake in the Biotechnologically Relevant *Bacillus megaterium*. Bioinformatics and biology insights, 6, 275.

Mbah, A. N., Mahmud, O., Awofolu, O. R., & Isokpehi, R. D. (2013). Inferences on the biochemical and environmental regulation of universal stress proteins from Schistosomiasis parasites. Advances and applications in bioinformatics and chemistry: AABC, 6, 15.

Sahoo, P. K., & Kim, K. (2013). A review of the arsenic concentration in paddy rice from the perspective of geoscience. Geosciences Journal, 1-16.

The article Johnson et al. 2010 was one of the selections by the publishers in Chapter 2 (Environmental Health) of a book titled “Issues in Environment, Health, and Pollution: 2011 Edition”



CHAPTER 2 ENVIRONMENTAL HEALTH

University of South Africa, Pretoria: The case for visual analytics of arsenic concentrations in foods

New research, ‘The case for visual analytics of arsenic concentrations in foods,’ is the subject of a report. “Arsenic is a naturally occurring toxic metal and its presence in food could be a potential risk to the health of both humans and animals. Prolonged ingestion of arsenic contaminated water may result in manifestations of toxicity in all systems of the body,” investigators in Pretoria, South Africa report.

“Visual Analytics is a multidisciplinary field that is defined as the science of analytical reasoning facilitated by interactive visual interfaces. The concentrations of arsenic vary in foods making it impractical and impossible to provide regulatory limit for each food. This review article presents a case for the use of visual analytics approaches to provide comparative assessment of arsenic in various foods. The topics covered include (i) metabolism of arsenic in the human body; (ii) arsenic concentrations in various foods; (ii) factors affecting ar-