



University of  
Massachusetts  
Amherst

## Developing a Toolkit for Citizen Scientists' Evaluation of Drinking Water Quality

Item Type	Master Projects
Authors	D'Andrea, LeighAnn
Download date	2026-03-11 17:05:41
Link to Item	<a href="https://hdl.handle.net/20.500.14394/4753">https://hdl.handle.net/20.500.14394/4753</a>

# Developing a Toolkit for Citizen Scientists' Evaluation of Drinking Water Quality

A Masters Project Presented

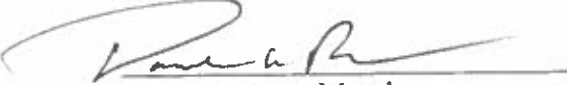
by

**LeighAnn D'Andrea**

Approved as to style and content by:



Emily Kumpel, Chairperson



Dave Reckhow, Member



Chul Park  
Civil and Environmental Engineering Department

## Acknowledgements

I'd like to thank my advisor, Dr. Emily Kumpel, for her support, guidance, and patience with me during my graduate studies at UMass. I was lucky enough to have the opportunity to work as her Teaching Assistant and Research Assistant. Her confidence in my abilities to do both kept me motivated to succeed. I would also like to thank Dr. Dave Reckhow for being on my committee and providing valuable comments to my studies and research.

## Abstract

This project evaluated tools for citizen science by assessing the potential to monitor drinking water quality at home with commercially sold, or off the shelf, test kits. The first objective was to compare the results of off the shelf test kits with laboratory-obtained values as performed in a controlled environment by trained laboratory personnel. The second objective was to provide potential citizen scientists with a selection of test kits to determine if they could successfully use the kits. These two objectives contributed to the end goal of understanding whether test kits on the market existed that could be used by citizen scientists to monitor drinking water quality and identify gaps and challenges in existing testing methods available to the public.

The drinking water parameters considered were iron, copper, manganese, and fluoride using combination test kits, which could measure multiple drinking water parameters in a single test, and single parameter test kits. These kits were evaluated for their ability to measure known concentrations of these four parameters in three water sources: DI water to represent water of high purity, Amherst tap water to represent a local water source, and water from the Mill River to represent a local surface water source. Typically, the kits measuring iron and copper performed best in the DI water matrix where there was nothing to interfere with the added constituent, and often performed poorly in Amherst tap water and Mill River water. The combination kits usually performed worse than the single parameter test kits. The evaluated kits measuring fluoride were unable to accurately measure fluoride concentrations in any water matrix, while the kit measuring manganese was able to reasonably approximate manganese concentrations in Amherst tap water and Mill River water. Overall, for most parameters, some kits worked acceptably well, and others did not at all, allowing guidance on selection of methods depending on the goal of a citizen science program. Overall, the test kits could benefit from more precise instructions to users and a method to consistently read the test kits' results between users.

## TABLE OF CONTENTS

	Page
ACKNOWLEDGMENTS.....	2
ABSTRACT.....	3
LIST OF TABLES.....	6
LIST OF FIGURES.....	7
CHAPTER	
1. DRINKING WATER QUALITY IN THE UNITED STATES.....	8
1.1. Introduction.....	8
1.2. Drinking Water Quality .....	8
1.2.1. Safe Drinking Water Act and Testing Drinking Water Quality.....	8
1.2.2. Water Quality Testing Available to the Public.....	9
1.2.3. Iron, copper, manganese, and fluoride in drinking water.....	9
1.3. Citizen Science.....	11
1.3.1. Levels of Citizen Science.....	13
1.3.2. Challenges and Benefits of Citizen Science.....	13
1.3.3. Citizen Science and Water Data.....	14
1.4. Research Objectives.....	16
2. COMPARING THE ACCURACY OF OFF-THE-SHELF TEST KITS UNDER LABORATORY CONDITIONS.....	17
2.1. Overview.....	17
2.2. Research Objectives.....	17
2.3. Background.....	18
2.4. Methods.....	25
2.4.1. Water Matrices.....	25
2.4.2. Stock Solutions.....	26
2.4.3. Experimental Methods.....	30
2.4.4. Statistical Analysis.....	31
2.5. Results.....	31
2.5.1. Iron.....	32
2.5.2. Copper.....	39

2.5.3. Fluoride.....	42
2.5.4. Manganese.....	43
2.6. Sensitivity Analysis.....	43
2.6.1. Iron.....	45
2.6.2. Copper.....	46
2.6.3. Fluoride.....	47
2.6.4. Manganese.....	48
2.6.5. Influence of pH on results.....	48
2.7. Ranking.....	49
2.8. Discussion.....	50
3. STUDENT GROUP CASE STUDY.....	54
3.1. Overview.....	54
3.2. Research Objectives.....	54
3.3. Methods.....	54
3.3.1. Study Design.....	54
3.3.2. Group Selection and Data Collection.....	55
3.4. Results.....	55
3.5. Discussion.....	59
4. CONCLUSION.....	60
REFERENCES.....	61
APPENDIX.....	67

## List of Tables

Table 1 Summary of Challenges and Associated Recommendations .....	13
Table 2 Characteristics of the evaluated off-the-shelf test kits .....	20
Table 3 Combination test kits and information on tested parameters .....	22
Table 4 Range and increments of OTS iron kits.....	23
Table 5 Range and increments of OTS copper kits .....	24
Table 6 Range and increments of OTS manganese kits.....	24
Table 7 Range and increments of OTS fluoride kits.....	25
Table 8 Target and measured concentrations of iron solutions.....	27
Table 9 Target and measured concentrations of copper solutions .....	28
Table 10 Target and measured concentrations of manganese solutions .....	29
Table 11 Target and measured concentrations of fluoride solutions .....	30
Table 12 Final Ranking of test kit brands based on parameter .....	50
Table 13 Color codes and lab-obtained solution concentrations of water samples brought to meetings with EWB and SWE for citizen science testing. ....	55
Table 14 Iron concentrations measured by test kits in DI Water. All tests were repeated five times. ....	67
Table 15 Iron concentrations measured by test kits in Amherst Tap Water. All tests were repeated five times .....	68
Table 16 Iron concentrations measured by test kits in Mill River Water. All tests were repeated five times .....	69
Table 17 Copper concentrations measured by test kits in DI Water. All tests were repeated five times.....	71
Table 18 Copper concentrations measured by test kits in Amherst, MA Tap Water. All tests were repeated five times .....	72
Table 19 Copper concentrations measured by test kits in Mill River Water .....	72
Table 20 Fluoride concentrations measured by test kits in DI Water .....	73
Table 21 Fluoride concentrations measured by test kits in Amherst Tap Water .....	74
Table 22 Fluoride concentrations measured by test kits in Mill River Water.....	75
Table 23 Manganese concentrations measured by test kits in DI Water .....	75
Table 24 Manganese concentrations measured by test kits in Amherst Tap Water.....	76
Table 25 Manganese concentrations measured by test kits in Mill River Water.....	76

## List of Figures

Figure 1 Levels of Citizen Science (adapted from Haklay, 2013).....	6
Figure 2 Iron (0-12 mg/L) as measured by combination test kit compared to the laboratory-obtained concentrations (FerroVer Iron Reagent).....	32
Figure 3 Iron (0-12 mg/L) as measured by the iron only test kit compared to the laboratory-obtained concentrations (FerroVer Iron Reagent).....	34
Figure 4 High range iron (20-100 mg/L) as measured by the iron only test kit and combination test kit compared to the laboratory-obtained concentration (FerroVer Iron Reagent).....	36
Figure 5 Copper concentrations (0-12 mg/L) as measured by combination test kit compared to the laboratory-obtained concentration (CuVer Iron Reagent).....	39
Figure 6 Copper concentrations (0-12 mg/L) as measured by copper only test kit compared to the laboratory-obtained concentration (CuVer Iron Reagent).....	41
Figure 7 Fluoride concentrations (0-12 mg/L) as measured by combination test kit compared to the laboratory-obtained concentration (SPADNS 2 Method).....	42
Figure 8 Manganese concentrations (0-2 mg/L) as measured by the manganese only test kit compared to the laboratory-obtained concentration (PAN LR Method).....	43
Figure 9 Definitions for sensitivity calculation based on detection.....	44
Figure 10 Definitions for sensitivity calculation based on the USEPA’s SMCLs.....	44
Figure 11 Sensitivity analysis of kits measuring iron based on ability to detect.....	45
Figure 12 Sensitivity analysis of kits measuring iron based on EPA’s SMCL of 0.3 mg/L of iron.....	45
Figure 13 Sensitivity analysis of kits measuring copper based on ability to detect.....	46
Figure 14 Sensitivity analysis of kits measuring copper based on EPA’s SMCL of 1.0 mg/L of copper.....	46
Figure 15 Sensitivity analysis of kits measuring fluoride based on ability to detect.....	47
Figure 16 Sensitivity analysis of kits measuring fluoride based on EPA’s SMCL of 2.0 mg/L of fluoride.....	47
Figure 17 Sensitivity analysis of kit measuring manganese based on ability to detect.....	48
Figure 18 Sensitivity analysis of kits measuring manganese based on EPA’s SMCL of 0.05 mg/L of manganese.....	48
Figure 19 Iron concentrations (0-100 mg/L) in the Amherst tap water matrix measured by the FerroVer Iron Reagent compared to the recorded pH.....	49
Figure 20 Iron concentrations (0-12 mg/L) as measured individually by the citizen scientists in the DI water matrix.....	56
Figure 21 Iron concentrations (0-12 mg/L) as measured individually by the citizen scientists in the Amherst tap water matrix.....	56
Figure 22 Iron concentrations (0 -12 mg/L) as measured in groups by the citizen scientists in the DI water matrix.....	57
Figure 23 Iron concentrations (0 -12 mg/L) as measured in groups by the citizen scientists in the Amherst tap water matrix.....	57
Figure 24 Average results from surveys.....	58

## Chapter 1: Drinking Water Quality in the United States

### 1.1. Introduction

Whether or not a person can go to their tap water and drink the water knowing they will not get sick is a key distinction between least-developed countries and high-income countries (Baum, Bartram, & Hrudey, 2016). However, improved drinking water supplies in affluent nations, such as the United States, can still be contaminated (Sobsey, Stauber, Casanova, Brown, & Elliott, 2008). After the water crisis that took place in Flint, Michigan, causing a declared state of emergency, polls have shown that more than half of Americans are not confident in the safety of their tap water (Baum et al., 2016; Pieper, Tang, & Edwards, 2017).

To better understand built and natural systems, monitoring programs need available and high-quality data (Greenwood, Mills, & Roig, 2007). However, there are numerous constraints preventing scientists and researchers from obtaining the necessary data. Therefore, it may be beneficial for the scientific community to empower and recognize the potential role of the citizen scientist. The process of citizen science involves involving the non-scientific community in some level of scientific research. Through proper use of citizens science, data could become increasingly available and used to better manage resources that contribute to access to safe drinking water in the future.

### 1.2. Drinking Water Quality

#### *1.2.1. Safe Drinking Water Act and Testing Drinking Water Quality*

The Safe Drinking Water Act (SDWA) was passed by Congress in 1974 with a mission to protect public health. The SDWA applies to every public water system in the United States. The SDWA delivers a framework for ensuring provision of safe drinking water to the public. The United States Environmental Protection Agency (US EPA) and states can take legal actions against public water systems that are not compliant with the SDWA. Community water systems are defined as any entity or institution that supplies water to an average of a minimum of 25 people or supplies water to at least 15 service connections for at least 60 days per year. Systems that do not meet this definition are considered private (or homeowner) water suppliers (EPA, 2004).

National standards for the types and levels of drinking water constituents are set based on reliable scientific evidence, while also considering financial expenses and available technologies.

The US EPA takes three steps when deciding on standards: identify the contaminant for further study, determine a maximum contaminant level goal to regulate, and determine a maximum contaminant level to be delivered to users of a public water system.

Monitoring the water quality in systems that meet the definition of community water systems is the responsibility of the systems themselves as well as state-level agencies, in accordance with the SDWA. Homeowners with private sources (i.e., wells) are responsible for testing themselves. Mass DEP recommends that well owners send samples to a state-certified lab, and test for arsenic, chloride, copper, fluoride, gross alpha screen (bedrock wells only), hardness, iron, lead, manganese, pH, radon, sodium, and VOCs initially and then every 10 years, and coliform bacteria and nitrate/nitrite initially and then at least every year; unless a local Board of Health requires more often (MassDEP, 2019a, 2019b)

### *1.2.2. Water Quality Testing Available to the Public*

The Mass DEP, and most other state agencies, recommend that if a homeowner wants to test their drinking water – whether from a public water supplier or a private well – that they use a state-certified lab, with fees set by the individual labs. However, there are numerous do-it-yourself test kits available for private citizens to purchase online or at local big box or hardware stores that citizens looking for a more affordable option may elect to do (MassDEP, 2018).

These do-it-yourself kits that are commercially sold do not undergo any formal certification or accreditation process to ensure their accuracy. However, formally certifying these test kits could fit into the scope of the National Sanitation Foundation (NSF). NSF provides accreditations to third parties to ensure their compliance with national and international standards. For instance, one service NSF provides is product certification of bottled water (NSF, 2019b, 2019a).

### *1.2.3. Iron, copper, manganese, and fluoride in drinking water*

The SDWA establishes National Primary Drinking Water Regulations, which set federally enforceable maximum contaminant levels (MCL) for parameters, and National Secondary Drinking Water Regulations, which set secondary maximum contaminant levels to serve as guidance for public water systems (EPA, 2004). While there are many possible contaminants regulated under these Drinking Water Regulations, we focus on iron, copper,

manganese, and fluoride. Some of these parameters are regulated by an MCL and SMCL. Also, these four parameters represent a range of advertised abilities of off the shelf test kits.

Iron is an element that occurs naturally on Earth and can enter bodies of water when it rains. The rainwater percolates through soil and rocks, which brings trace amounts of iron into bodies of water. This water is then used later for various human purposes such as agriculture, recreation, and drinking (Khatri, Tyagi, & Rawtani, 2017). Iron can also occur in drinking water systems from corrosion of iron pipes or the residual of iron-based coagulants. The challenge is that iron exists in drinking water in various forms. Iron exists in water as either soluble, able to dissolve in water, or insoluble. When classifying types of iron in terms of solubility it is important to consider the pH of the sample: the lower the pH level of the water, the more soluble the iron will be (Metzger, 2005). The soluble, reduced form of iron is ferrous iron (iron (II) or  $\text{Fe}^{2+}$ ). When ferrous iron is present in drinking water there are no physical changes. The insoluble, oxidized form of iron is ferric iron (iron(III) or  $\text{Fe}^{3+}$ ). Total iron refers to the combination of ferrous and ferric iron. Once a water sample containing ferrous iron is exposed to air it oxidizes to ferric iron. An important parameter that affects this redox reaction is the pH. As pH increases the conversion of ferrous to ferric iron increases. Ferrous iron can serve as a substrate for various types of iron oxidizing bacteria; bacteria can become incorporated into biofilms on pipe walls and use the ferrous iron and oxidize it to ferric iron (Khatri et al., 2017). Iron bacteria growth is dependent on pH, the range for bacteria growth being 5.5 to 8.2 where 6.5 is the optimum pH (Metzger, 2005).

The SMCL for iron set by the EPA is 0.3 mg/L. When iron is present in drinking water above this concentration, consumers may notice a difference in their drinking water, such as aesthetic and technical effects. The EPA notes the noticeable changes for iron above the SMCL include a rusty color, sediment (loose deposits in the distribution system), metallic taste, and a reddish color.

Copper is a naturally occurring metal in rocks, soils, plants, animals, and water. Copper can be found in various types of water including surface water, groundwater, seawater, and drinking water, however it is typically in complexes or as particulate matter. It is also commonly used as a household plumbing material. In drinking water, copper concentrations can vary with water characteristics such as pH or availability of copper in the distribution system. When water is running, copper concentrations are typically lower. In stagnant waters, concentrations are

likely to be higher due to leaching from pipe walls. In groundwater, copper concentrations are generally low, therefore copper typically enters drinking water through direct contamination of well water or corrosion of copper pipes from acidic water (CDC, 2018). Copper is regulated as an MCL and SMCL with concentrations of 1.3 mg/L and 1.0 mg/L respectively. Regulation of the MCL for copper falls under the Lead and Copper Rule which has a purpose of minimizing lead and copper levels in drinking water by reducing the water corrosivity. At concentrations above the MCL consumers are at risk of adverse health effects such as stomach distress and liver or kidney damage (EPA, 2008)

Manganese naturally occurs in some surface waters and groundwaters. Various human activities can cause manganese contamination in drinking water supplies, although the main avenue for human consumption of manganese is through food. In groundwater manganese exists primarily as  $Mn^{2+}$  due to lack of oxygen, but in subsurface waters it may exist in a variety of oxidation states as soluble complexes or suspended particles (WHO, 2011). Manganese is regulated by the EPA with an SMCL of 0.05 mg/L. At concentrations above this, consumers will notice aesthetic changes but will not be at risk for adverse health effects (EPA, 2004).

Fluoride is a special case because it is of significant global concern. Parts of the world, most notably India and China, are suffering from fluorosis, due to the contamination of their groundwater with high concentrations of fluoride (Farooqi, Masuda, & Firdous, 2007). The element fluorine does not exist in nature due to its high reactivity and therefore in water is only a concern as fluoride. In the United States, fluoride is added to water in low concentrations in order to strengthen teeth and hopefully offset dental costs (WHO, 2004).

### 1.3. Citizen Science

People inhabiting the planet have been informally writing down their observations of the natural world around them for centuries (Miller-Rushing, Primack, & Bonney, 2012). However, this idea of implementing and accepting a formal process of having citizen scientists - which involves getting the non-scientific, non-academic community involved in some level of scientific research - is relatively new to the scientific community. Community based research, public participation, community-based monitoring, participatory science, partnership approaches, and many other terms relate to the fundamental concept of citizen science. Having the public involved in studies that affect their lives can lead to important new insight into their health and

environment; too often, the community is a discounted part of an ecosystem (C. T. Conrad & Daoust, 2008). Considering and using the community through citizen science could broaden abilities of researchers and lead to new, previously unattainable, scientific insight (Miller-Rushing et al., 2012).

Citizen science is likely to be popular in the coming years and is not currently being used to its full potential (Barbara A. Israel, Amy J. Schulz, Edith A. Parker, & Becker, 1998; Bonney et al., 2009). There are two gaps of information causing a hindrance to the growth of citizen science. The first is that there is insufficient analysis of successful citizen science programs; better understanding similarities in what constitutes a successful citizen science testing program would help inform other similar programs trying to achieve success. The second gap is that there are not enough case studies showing citizen science data being used by decision makers (C. C. Conrad & Hilchey, 2011).

While many reasons may account for the recent surge in citizen science in the past few decades, one likely explanation is advances in technology that enable instantaneous communication and analysis, which aids the process of citizen science. Another contributing factor is professionals realizing that citizen science can provide free labor in their research projects (Silvertown, 2009). Decreases in government funding for environmental and monitoring programs is incentivizing citizens to take it upon themselves (Au et al., 2000). Access to data is key in environmental protection (Ottinger, 2017). In addition, the public's general mistrust in their governments' concern and care for the environment is also contributing to the increase in citizen science projects and participation (Au et al., 2000). There is also a desire to use citizen science programs to educate people – and particularly middle and high school students - by partnering with established organizations (Rawot, Peirce, & Kanoy, 2011). This is effective because it promotes both understanding as well as appreciating science by actively participating in it (Silvertown, 2009).

### 1.3.1. Levels of Citizen Science

One method to classify the levels of citizen science is to do so by the level of engagement and participation required on the part of the citizen scientist. As the required engagement from the citizen scientist increases, so does the level. Literature has divided this into four levels illustrated in Figure 1 (Assumpção, Popescu, Jonoski, & Solomatine, 2018).

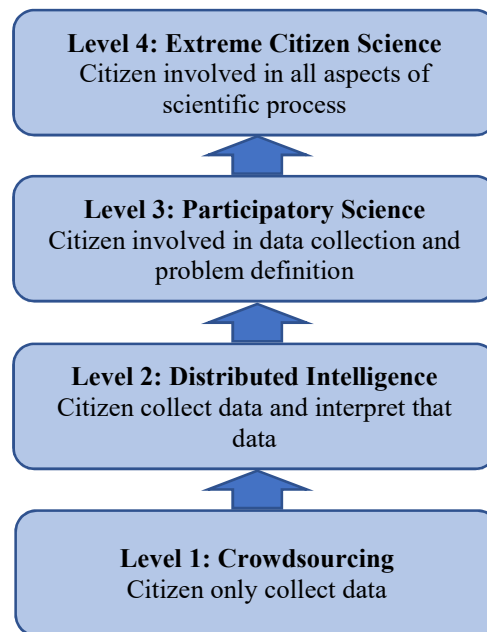


Figure 1 Levels of Citizen Science (adapted from Haklay, 2013)

### 1.3.2. Challenges and Benefits of Citizen Science

Citizens science programs, as any other program, include benefits and challenges that are important to be aware of and consider before making the decision to utilize this process. On the positive side, it is argued that consulting the public results in overall better environmental planning and an increased support from the public on government actions and policy since they are presumably more involved in the process that caused the actions (Au et al., 2000). Citizen science is important to public policy because it is inclusive of the affected population from the start (Dickinson et al., 2012). The key challenges facing citizen science and recommendations for possible solutions are outlined in Table 1 (Barbara A. Israel et al., 1998; C. T. Conrad & Daoust, 2008; Newman, Graham, Crall, & Laituri, 2011; Sullivan et al., 2014).

**Table 1.** Summary of Challenges and Associated Recommendations

Challenge	Description	Recommendation
Achieving trust and mutual respect	The quality of citizen data is not trusted by the researcher. The intention of the researcher/partner organization is not trusted by the citizens	Jointly develop operational norms for the program. Be flexible in methodology and have unique criterion for each data set
Properly distributing power and control	Making sure power is distributed effectively to public and researcher	Democratic leadership
Different perspectives on priorities, goals, values, etc.	Different perspectives from public vs. researcher or within the same community	Clearly define program mandates
Funding	Hard to obtain funding, but also hard to distribute amongst partners once obtained	Jointly develop operational norms

Time	What portion of the scientific process should be allocated what time. Also, time to build trust	Create teams with different responsibilities. Interdependence not independence.
Representation	Who is representing the community? Are necessary people being left out of key meetings?	Presence of community organizer who is already a leader
Proving Success	Mutually agree upon program success	Jointly define goals and objectives at the start of the program. Formal evaluation.
Lack of policy	No issued policy or required framework on citizen science	Clearly define program mandates
Quality vs. Quantity	Achieving balance of quality and quantity. Achieving reliability	Assessment of citizens at start of program.
Data Management and Analysis	Programs are not equipped to handle and analyze large amounts of data	Cyber-infrastructure support systems

Table 1 shows that many of the challenges that citizen science faces are within the program itself. Maintaining the relationship between the citizen and the researcher is crucial for success.

However, even considering these challenges, citizens science is gaining recognition in the scientific community and being praised for its ability to not only collect data but also build public engagement. Yet, citizen science is still rarely used or accepted as a primary tool in scientific research literature. Burgess et al., conducted a survey to decipher why this trend exists. 423 scientists and 125 citizen science program managers were questioned, and the survey revealed four major trends. The trends were; lack of awareness amongst scientists of citizens science projects that match their research needs, lack of confidence amongst scientists their research is well-suited for citizen science, inconsistency in data quality across projects, and concern for bias. To get citizen science more recognition and respect it was suggested that programs be more transparent, specifically regarding their data and methods (Burgess et al., 2017).

### *1.3.3. Citizen Science and Water Data*

A citizen science monitoring program has its own specific set of objectives. These objectives may include but are not limited to management, awareness, education, serendipity, recreation, social research, economic research, knowledge, improvement (Tulloch, Possingham, Joseph, Szabo, & Martin, 2013). Citizen science monitoring can have both scientific and social objectives, which is one of the advantages of using citizen science. For monitoring programs, large sets of data are required and imperative to success (Bonney et al., 2014). These large quantities of data are also necessary because citizens are typically only taking one measurement; however, these data should also be of high quality. It is impossible to get high quality data without high quality

data input and management. Striking a proper balance between the data quality and quantity is crucial to success but has proven to be a difficulty. Furthermore, identifying and assessing the ability of the citizens will help the program organizers decide what sort of data they can obtain accurately (Sullivan et al., 2014). Two common challenges that citizen science monitoring programs face regarding data are accounting for nonstationary and controlling for sources of variation (Hochachka et al., 2012). Many studies agree that collecting meta-data on aspects of environment and survey execution will prove to be useful and help to standardize results (Bird et al., 2014). It is important to recognize these challenges prior to creating a monitoring program so it can be accounted for during program design. Throughout the process, the organization of a monitoring program will decide its overall success.

Being organized maximizes the impact of a program and makes it more likely that participants will be able to collect such high amounts of data. This is a sound example of citizen science not reaching its full potential. For instance, with enough organization, citizen science monitoring could be activated in the event of a natural disaster and collect valuable data. To achieve this high level there are resources available to help citizen science programs. The Citizen Science Association is an organization with a website set up specifically to support citizen science programs with data management, scientific precision, and overall project evaluation (Bonney et al., 2014). Although the Citizen Science Association is the most notable example, it is not the only resource out there for hopeful programs. It is also important to consider that resources such as this not only support individual programs, but they also create a hub of information for other programs to analyze and learn from.

Citizen science monitoring programs can be used to specifically monitor water quality, and it has proven to be effective (C. T. Conrad & Daoust, 2008). Consistent water monitoring is the most effective way to protect the public from waterborne pathogen exposure, and citizen science can be a tool for achieving this consistency (Farnham et al., 2017). Especially in a monitoring program for water quality, it is important to define standards early in the process. Standards create an indisputable common ground for the researchers and citizens involved. The standards help to decipher the reliable from the unreliable and the relevant from the irrelevant (Ottinger, 2010). However, it is recognized that citizen science is a relatively new process, therefore it would be effective to start with lower standards and raise them over time as it becomes a more established process (Sullivan et al., 2014).

#### 1.4. Research Objectives

This research aims to develop a toolkit for citizen science to use to monitor drinking water quality. To achieve this aim, we evaluated whether currently available, off-the-shelf test kits can accurately measure several water quality parameters (copper, iron, fluoride, and manganese). First, we measured whether nationally available at-home water quality testing kits could accurately measure concentrations of these parameters under controlled laboratory conditions (Chapter 2). Second, we provided citizen scientists with a selection of the test kits to see if they could successfully deploy them while also getting feedback on their experiences using the test kits (Chapter 3). The end goal was to produce a ranking of off the shelf test kit that could be suited for citizen scientists' monitoring of drinking water quality; to inform the public about the efficacy of these available methods; and to identify whether there are gaps in available test methods.

## **Chapter 2: Comparing the of accuracy of off-the-shelf test kits under laboratory conditions**

### 2.1. Overview

Off-the shelf (OTs) test kits were purchased online from a major online retailer (www.amazon.com), just as a citizen scientist would obtain a test kit. While consumers might also purchase test kits from local stores (e.g., big box stores, local hardware stores), we wanted to get kits that would be nationally available and so selected through the online domain. Kits were selected based primarily on popularity and price, as a consumer might decide to purchase a kit. A variety of test kit brands were selected including some that measure for one drinking water constituent and some that measure multiple constituents in a single test. A kit that measured for multiple constituents in a single test was referred to as a combination kit.

The accuracy of the off-the-shelf test kits was examined in the lab for four common drinking water contaminants: iron, copper, manganese, and fluoride. The experiments were repeated for three sources of water: DI water, Amherst tap water, and Mill River water. For each contaminant, a stock solution of high concentration of the constituent was made. This stock solution was then diluted, using the appropriate source water, in order to achieve a range of lower concentrations. The range of lower concentrations for analysis was selected based on what the test kits were marketed to measure and any advisory or action levels set by the EPA. The concentrations of the diluted solutions were tested with the DR2700 Hach Spectrophotometer, which was taken to be the true, laboratory-obtained concentration of the solution. Each solution was tested with each brand's test in quintuplicate. Each constituent was tested with at least the four combination test kits. Iron, copper, and manganese were tested using additional kits that measured for only that constituent.

### 2.2. Research Objectives

The objectives of the laboratory portion of this research project were to:

- Compare the accuracy of measurements obtained from off-the-shelf tests kits for measure iron, copper, manganese, and fluoride in drinking water quality with laboratory-obtained concentration; and
- Produce a ranking of test kits according to accuracy.

### 2.3. Background

Four drinking water contaminants were selected for analysis to achieve the stated research objectives: iron, copper, manganese, and fluoride. These constituents are regulated by the EPA under either primary or secondary standards. Primary standards are legally enforceable and are set to protect public health, while secondary standards help to control cosmetic, aesthetic, and technical effects. Copper is regulated by primary standard with a maximum contaminant level (MCL) (US EPA, 2015b). Iron and manganese are regulated as secondary standards, which means these secondary maximum contaminant level (SMCL) concentrations are guideline values which are not federally enforceable. These secondary standards serve as recommendations, not requirements, for systems so they can successfully control cosmetic, aesthetic, and technical effects of their drinking water (US EPA, 2015a). Fluoride is considered both a primary and secondary contaminant: the MCL is 4.0 mg/L and the SMCL is 2.0 mg/L, for which the EPA does require special notification if a water system exceeds the secondary standard for fluoride. This is the only drinking water constituent that requires any notification when exceeding a secondary standard.

**Iron.** The SMCL for iron in drinking water is 0.3 mg/L; greater concentrations in water do not pose a health risk but can negatively affect aesthetics or taste of water and cause infrastructure damage (US EPA, 2015a). The evaluated combination test kits measured iron up to a much higher concentration than the test kits measuring only iron: each combination test kit claimed it was capable of measuring iron in water up to a maximum detection level of 500 mg/L. Three of the kits measuring only iron claimed to measure up to 5 mg/L, and one test kit measuring up to 100 mg/L. While the combination kits measured higher concentrations, they had larger intervals between measurement points. For instance, a combination kit may measure 0, 5, and 10 mg/L of iron, while a test measuring only iron might measure 0, 0.02, 0.05, 0.1, 0.2, and 0.3 mg/L of iron. In drinking water, the primary concern is the SMCL of 0.3 mg/L of iron; therefore, it is important for our purposes that a test kit can measure near that concentration.

**Copper.** Copper in drinking water systems is regulated by the EPA under the Lead and Copper Rule. The purpose of this rule, according to the EPA, is to protect public health by minimizing lead and copper levels in drinking water, largely through reducing water corrosivity. This is because the primary way lead and copper enter drinking water is from the corrosion of lead and copper plumbing materials. Therefore, if there is copper present it is possible that lead is

present or vice versa (EPA, 2008). The EPA also has an SMCL of 1.0 mg/L for copper in drinking water. Concentrations greater than this will cause a metallic taste and a blue staining (US EPA, 2015a). All but one evaluated test kit measured copper concentrations up to 10 mg/L; one combination test kit measured copper up to 300 mg/L.

**Manganese.** Manganese in drinking water is regulated by the EPA with an SMCL of 0.05 mg/L in drinking water. At concentrations greater than this, consumers will notice a black or brown color, black staining, and a bitter taste (US EPA, 2015a). None of the evaluated combination test kits measured for manganese in drinking water. Only one test kit that measured manganese concentration only was available via the online retailer.

**Fluoride.** The EPA regulates fluoride in drinking water as an MCL and an SMCL. The SMCL is 2.0 mg/L for fluoride in drinking water. At concentrations between 2.0 – 4.0 mg/L, there are not expected to be adverse health effects, however, consumers with prolonged exposure may notice tooth discoloration. The MCL is 4.0 mg/L for fluoride in drinking water. At concentrations greater 4.0 mg/L, there can be adverse health effects, including fluorosis (bone disease). Fluoride is commonly added to drinking water supplies to promote strong teeth and prevent cavities. It can also occur naturally in groundwater or enter water supplies from the discharge of various fertilizers and factories (US EPA, 2015a). Some parts of the world outside of the United States have groundwaters with severe fluoride contamination. This fluorosis causes serious adverse health effects for their consumers such as molted teeth and bone disease (Farooqi et al., 2007). No test kit that measured only the concentration of fluoride was available via the online retailer and therefore not evaluated in this research. The four combination test kits evaluated in this study measure the fluoride concentration of a water sample.

**Evaluated off-the-shelf test kits.** Four combination kits, four iron kits, two copper kits, one manganese kit, and one fluoride kit were evaluated (Table 2). Among the 11 tests chosen for this research project, only 4 provided guidance on the amount of water to use for each test and none provided guidance on whether or how to flush (or not flush) water first or the time of day to take the water sample (Table 2). These are two considerations (out of many) that could potentially effect results of drinking water quality parameters.

Table 2. Characteristics of the evaluated off-the-shelf test kits.

<b>Combination Test Kits</b>							
<b>Test Kit Brand</b>	<b>Measurement</b>	<b>Cost per container</b>	<b>Tests per container</b>	<b>Cost per test</b>	<b>Guidance on sample volume?</b>	<b>Guidance on flushing or sampling time?</b>	<b>Measures at relevant SMCL or MCL?</b>
Baldwin Meadows 10in1	Copper, fluoride, free chlorine, iron, lead, nitrate, nitrite, pH, total alkalinity, total hardness	\$18.99	100	\$0.19	No	No	No
JNW 10in1	Copper, fluoride, free chlorine, iron, lead, nitrate, nitrite, pH, total alkalinity, total hardness	\$19.99	125	\$0.16	No	No	No
Swiss Labs 14in1	Alkalinity, bromine, carbonate, copper, cyanuric acid, fluoride, free chlorine, hardness, iron, lead, nitrate, nitrite, pH, total chlorine	\$19.95	150	\$0.13	No	No	No
Technically Pure 14in1	Bromine, carbonate, copper, cyanuric acid, fluoride, free chlorine, iron, lead, nitrate, nitrite, pH, total alkalinity, total chlorine, total hardness	\$18.99	100	\$0.19	Yes	No	No
<b>Iron Test Kits</b>							
Aquacheck Hach Iron Strips	iron	\$25.25	25	\$1.01	Yes	No	Yes
Bartovation	iron	\$18.88	50	\$0.38	No	No	No
Sensafe Iron Check	iron	\$19.99	25	\$0.80	No	No	Yes
WaterWorks	iron	\$26.54	50	\$0.53	Yes	No	Yes
<b>Copper Test Kits</b>							
Aquacheck Hach Copper Strips	copper	\$12.94	25	\$0.52	No	No	No
Sensafe Copper Check	copper	\$13.29	25	\$0.53	Yes	No	No
<b>Manganese Test Kits</b>							
Sensafe Manganese Check	manganese	\$28.31	24	\$1.18	Yes	No	Yes

**Combination Kits.** Four combination test kits were used in this project; Baldwin Meadows 10in1, JNW 10in1, Swiss Labs 14in1, and Technically Pure 14in1. These four brands used strips; these strips are dipped in the water sample and then the color appearing on the strip is matched to a color chart provided on the test kit bottle. The combination kits capable of testing multiple parameters cost between \$0.13-\$0.19 per test (with each test evaluating for 10-14 parameters within ‘one test’); all of these were much cheaper than the single parameter-only test. The Swiss Labs 14in1 kit yielded the cheapest cost per test strip of \$0.13 (Table 2). The Technically Pure 14in1 kit was the only combination kit that provided guidance on how much water to use as a sample volume: this method included four sample 10 mL test tubes made of plastic in the box, and the directions instructed users to fill the test tube to the top. All four of these kits used the words easy, accurate, and fast or quick on the bottle and in the instructions to market their product. All four test kits measured the following nine parameters; copper, iron, nitrate, nitrite, free chlorine, lead, pH, total alkalinity, and total hardness. In addition to these nine parameters, the JNW10in1 measured fluoride while the Swiss Labs14in1 and Technically Pure 14in1 measured fluoride, total chlorine, bromine, cyanuric acid, and carbonate in addition to those nine parameters. None of the combination kits were able to measure manganese.

The instructions for all the combination kits told users to dip a strip into a water sample for a set time, remove it while shaking off excess water, and then wait between 15 - 60 seconds. After that, the user was instructed to match the colors on the test strips to a color chart on the bottle. The Swiss Labs 14in1 kit specified exactly how long to wait before matching colors for each individual parameter. The other test kits had a single wait time for all parameters.

The kits offered varying levels of information to users on interpreting their results (Table 3). The JNW10in1 and Swiss Labs14in1 had boxes around the ‘ideal range’ or ‘target level’ respectively for each parameter. The Technically Pure 14in1 has the ‘optimal level’ for each parameter written in green versus written in black for all other levels. The Baldwin Meadows10in1 uses words such as; “OK”, “high”, and “low” written above the concentration levels on the color chart in order to provide users with information about their results.

Table 3. Combination test kits and information on tested parameters

	MCL / SMCL (mg/L)	Baldwin Meadows 10in1 (ppm)	JNW 10in1 (ppm)	Swiss Labs 14in1 (mg/L)	Technically Pure 14in1 (mg/L)
<b>Iron</b>	0.3 <sup>a</sup>	0-0.3 = OK 0.5 – 5 = High	0 = ideal range	0 = target level	0 = optimal level
<b>Copper</b>	1.3 <sup>b</sup>	0-0.4 = OK 1-2 = High	0 – 0.5 = ideal range	0 – 1 = target level	0 = optimal level
<b>Fluoride</b>	2.0 <sup>a</sup> , 4.0 <sup>b</sup>	0-4 = OK 5 = High	0 = ideal range	No target level identified	1 = optimal level

The provided information is not consistent across the combination test kits for each drinking water constituent (Table 3). For instance, the Baldwin Meadows 10in1 tells users that 0 – 0.3 ppm is ‘OK’ for iron in drinking water while 0.5 – 5 ppm is ‘High’, the remaining combination test kits tell users that 0 ppm or mg/L is the ideal range for iron in drinking water. Looking at copper; the Baldwin Meadows 10in1 defines 1 – 2 ppm as ‘High’ but the Swiss Labs 14in1 has 1 mg/L included in its target level for copper in drinking water. The Baldwin Meadows 10in1 identifies 0 – 4 ppm as ‘OK’ for fluoride in drinking water, while the JNW 10in1 identifies 0 ppm as the ‘ideal range’. The Swiss Labs 14in1 does not identify a target level for fluoride and the Technically Pure 14in1 identifies 1 mg/L as the optimal level. Although these kits measure many of the same parameters, they are not consistent in what the results mean. The JNW 10in1 was the only kit that instructs users to ‘take action if needed’ on the bottle. The remaining kits refer users to a website or an e-book.

**Iron Kits.** In addition to the four combination kits, four test kits that measured solely the iron concentration of a water sample were used to test the iron solutions of known concentration. These were: Aquacheck Hach Iron Strips, Sensafe Iron Check, WaterWorks, and Bartovation. These kits cost ranged from \$0.38 - \$1.01 per test, with one test capable of testing one water sample for only its iron concentration. The Aquacheck Hach Iron Strips, Sensafe Iron Check, and Bartovation kits all involved strips that involved dipping a strip in water sample and matching a color on the strip to a color chart provided on the test kit bottle, while the WaterWorks kit had users match the color of the liquid in the test tube with colors provided on a color chart. Three of the kits provided a sample vial: the Aquacheck Hach Iron Strips kit provided users with a 100 mL bottle for their water sample, the Sensafe Iron Check provided a 5 mL bottle, and the

WaterWorks provided a 5 mL plastic test tube. Two of these kits also instructed users to add a reducing agent to the sample: the WaterWorks instructed users need to fill the test tube to a line with their sample, put in a packet of reagent and shake for 20 seconds, wait for 3 minutes, and then; the Aquacheck Hach Iron Strips also involved adding a reducing agent to the water sample.

The eight kits used to measure iron solutions in the lab had different measurement ranges and increments (Table 4). As previously stated, the combination kits sometimes had higher detection abilities and larger increments, while the kits measuring only iron measure up to lower concentrations at smaller increments. These increments and ranges determined the concentrations we used in the lab when making iron solutions.

**Table 4.** Range and increments of OTS iron kits

Test Kit	Measurement Range (mg/L)	Increments
Baldwin Meadows 10in1	0 - 5	0, 0.3, 0.5, 1, 3, 5, 10, 25, 50, 100
JNW 10in1	0 - 500	0, 5, 10, 25, 50, 100, 250, 500
Swiss Labs 14in1	0 - 500	0, 5, 10, 25, 50, 100, 250, 500
Technically Pure 14in1	0 - 500	0, 5, 10, 25, 50, 100, 250, 500
Aquacheck Hach Iron Strips	0 - 5	0, 0.15, 0.30, 0.60, 1, 2, 5
Sensafe Iron Check	0 - 5	0, 0.02, 0.05, 0.10, 0.20, 0.30, 0.50, 0.75, 1, 2, 5
WaterWorks	0 - 5	0, 0.3, 0.5, 1, 3, 5
Bartovation	0 - 100	0, 2, 5, 10, 25, 50, 100

**Copper Kits.** In addition to the four combination kits, two test kits that measured solely the copper concentrations of water samples were used to test copper solutions of known concentrations in the lab: the Aquacheck Hach Copper Strips and Sensafe Copper Check. These kits cost about \$0.50 each (the Aquacheck Hach Copper Strips cost \$0.52 per test and the Sensafe Copper Check was \$0.53 per test) (Table 2). One test measures only the copper concentration in one water sample. Neither of these kits provided guidance on flushing or time of day to take a sample (Table 2). The Aquacheck Hach Copper Strip did not instruct users on a sample volume, while the Sensafe Copper Check instructs users to test 200 mL of water (although there is no sample vial for users to measure 200 mL). Neither test measured the exact MCL set by the EPA of 1.3 mg/L, but the Sensafe Copper Check did indicate the value of the MCL. These kits both involved dipping a strip in water and matching a color to a provided color chart after a specified amount of time. The Aquacheck Hach Copper Strip instructs users to dip water into sample for 5 seconds while moving back and forth (or only 1 second if the sample is above 86°F), hold the strip parallel to the ground for 60 seconds, and then match it to the

provided color chart. The Sensafe Copper Check instructs users to dip the strip into the 200 mL sample for 15 seconds with a constant back and forth motions, shake strip upon removal of water sample to get rid of excess water, and then wait 30 seconds before matching to color chart. This kit also tells users that they must complete color matching within 60 seconds of removal from water sample.

The six kits used to measure copper concentration all had different measurement ranges and increments (Table 5).

**Table 5.** Range and increments of OTS copper kits

<b>Test Kit</b>	<b>Measurement Range (mg/L)</b>	<b>Increments</b>
Baldwin Meadows 10in1	0-10	0, 0.5, 1, 3, 10
JNW 10in1	0-10	0, 0.5, 1, 3, 10
Swiss Labs 14in1	0-10	0, 0.5, 1, 3, 10
Technically Pure 14in1	0-300	0, 1, 10, 30, 100, 300
Aquacheck Hach Copper Strips	0-3	0, 0.2, 0.5, 1, 3
Sensafe Copper Check	0-2	0, 0.05, 0.1, 0.2, 0.4, 1, 2

**Manganese Kits.** None of the four combination test kits were able to measure manganese in drinking water. The Sensafe Manganese Check, which measured solely manganese concentration in a single test, was used. This kit included 24 tests for \$28.31 which yielded a cost of \$1.18 per test. The Sensafe Manganese Check involved dipping test strips and matching them to a color chart. This test kit identified the following maximum levels on the color chart; EPA = 0.05 ppm, Bangladesh = 0.1 ppm, and WHO = 0.4 ppm. These three levels were written in red while all other concentration increments were written in black. Table 6 identifies the range and increments of the Sensafe Manganese Check.

**Table 6.** Range and increments of OTS manganese kits

<b>Test Kit</b>	<b>Measurement Range (mg/L)</b>	<b>Increments</b>
Sensafe Manganese Check	0-1.6	<0.02, 0.05, 0.1, 0.2, 0.4, 1.0, >1.6

This kit provides users with a 5mL plastic sample tube and three packets with different contents. The instructions include six steps split into three sample preparation steps and three-color development steps. First, users are told to fill the provided plastic sample vial with the water sample to the top line and then dip the strip that is in the packet labeled Mn Strip #1 into the vial for 20 seconds while moving it gently back and forth. The final sample preparation step is to dip the strip that is in the packet labeled Mn Strip #2 into the vial for 20 seconds while

moving it gently back and forth. Users are warned that Mn Strip #2 contains 40mg of cyanide and are to avoid contact with skin and eyes. The color development steps include dipping the strip in the packet label Mn Strip #3 in the vial for 30 seconds with a speed of approximately 2 strokes per second and, upon removing the strip, shaking it once, briskly, to remove the excess water. Lastly, the user is told to wait 3 minutes before matching the color on the strip to the color on the chart. The instructions note that this is a modified version of the 1-(2-pyridylazo)-2-naphthol (PAN) method. There is also a diagram that illustrates how users should fold the strip in order to have best color matching results. Users are also told to dispose of liquid sample and Mn Strip #2 ‘in accordance with local environmental laws’ due to the amount of cyanide in them.

**Fluoride Kits.** No test kits were available through the online retailer to measure solely fluoride concentration, therefore, only the four combination test kits were used. The ranges and increments that the combination kits measure for fluoride are presented in table x.

**Table 7.** Range and increments of OTS fluoride kits

Test Kit	Measurement Range (mg/L)	Increments
Baldwin Meadows 10in1	0-5	0, 0.5, 1, 2, 4, 5
JNW 10in1	0-100	0, 10, 25, 50, 100
Swiss Labs 14in1	0-100	0, 10, 25, 50, 100
Technically Pure 14in1	0-5	0, 0.5, 1, 2, 5

## 2.4. Methods

### 2.4.1. Water Matrices

All experiments were performed using three source waters: DI water, Amherst tap water, and Mill River water. DI water was used to represent a control water source of high purity, because there would be no background concentration of any contaminants in that water. Amherst tap water was used to represent local water that citizen scientists would have access to. The tap water is a surface water that was treated at the Amherst water treatment plant and traveled through the distribution system before experimentation. The Mill River water also represented local water that was accessible to citizen scientists and an untreated surface water.

We originally intended to use a phosphate buffer to hold the pH of each solution at a high and a low pH. However, the addition of the buffer interfered with the iron solutions, therefore, we recorded the pH and temperature for each solution but did not adjust either.

#### 2.4.2. Stock Solutions

**Iron.** A 10 mg/L stock solution of iron was made by adding 0.0251 g of Ferrous Sulfate Heptahydrate ( $\text{FeSO}_4 \cdot 7\text{H}_2\text{O}$ ) to 500 mL of DI water. This stock solution was then diluted with DI water to achieve target concentrations of 0.1, 0.3, 0.5, 1.0, 3.0, 4.0, and 5.0 mg/L. These increments were selected based on the EPA's SMCL of 0.3 mg/L for iron in drinking water, as well as to align with the increments that the test kits are marketed to measure (Table 4). To test the higher concentrations measured by the combination test kits, a second stock solution of 150 mg/L was made by adding 0.375 g of  $\text{FeSO}_4 \cdot 7\text{H}_2\text{O}$  to 500mL of DI water. This stock solution was then diluted with DI water to achieve target concentrations of 10, 25, 50, and 100 mg/L.

All concentrations were measured with FerroVer Iron Reagent Power Pillows from Hach following Method 8008, the USEPA FerroVer Method, adapted from the Standard Methods for the Examination of Water and Wastewater. This method could measure iron concentrations between 0.02 – 3.00 mg/L; samples with a concentration  $>3.0$  mg/L had to be diluted with DI water before it could be measured in the DR2800 Spectrophotometer (spec). Method 8008 involved starting Program 265 Iron, FerroVer on the spec. The sample was prepared by adding 10 mL of the water sample to the sample cell and adding in the contents of one FerroVer Iron Reagent Powder Pillows. The cell was then swirled to mix and a timer for 3 minutes was started. During those 3 minutes the blank was prepared, which consisted of adding 10 mL of the sample to the sample cell without adding a powder pillow. The blank was put into the spec and used to zero the program. After 3 minutes, the sample cell was ready to be read. Method 8008 converts all soluble and most insoluble iron to soluble ferrous iron. Method 8008 specifies that copper concentrations will not interfere with this method because there is a masking agent included in the reagent.

This procedure was repeated for the other two source waters after adjusting for any background concentration of iron that was already present in the water sample. For example, during one of the days, the measured iron concentration in the Amherst tap water was 0.04 mg/L; therefore, the stock solution added to the Amherst tap water was adjusted to meet the desired end concentration (e.g. 0.3 mg/L). The goal was to keep the target concentrations uniform across the three water source experiments. The target concentrations and measured concentrations were recorded (Table 8). Any target concentration that is repeated in Table 8 indicates that part of the experiment was done at a different time due to availability of the test kits. Any concentration  $>0$

mg/L listed in the row where the target concentration was 0 mg/L represents a background concentration present in that water source (Table 8). Solutions were made fresh immediately before experiments.

**Table 8.** Target and measured concentrations of iron solutions

Target Concentration (mg/L)	Measured Concentration using FerroVer Iron Reagent Power Pillows (mg/L)		
	DI Water	Amherst tap water	Mill River water
0.0	0.00	0.04	0.02
0.1	-	-	0.06
0.1	0.12	0.24	0.20
0.3	0.33	0.39	0.36
0.3	0.30	0.33	0.33
0.5	0.55	0.60	0.53
0.5	0.44	0.53	0.67
1.0	1.05	1.17	0.94
1.0	1.01	0.93	0.85
1.0	1.91	1.89	1.79
3.0	3.10	3.30	3.04
4.0	4.42	4.22	3.86
5.0	5.60	4.84	4.90
5.0	5.58	4.54	4.24
10.00	11.00	7.60	7.50
25.00	23.00	22.50	23.00
50.00	44.00	48.00	42.00
100.00	107.00	97.00	104.00

**Copper.** A 10 mg/L stock solution of copper was made by adding 0.0391 g of Cupric Sulfate (CuSO<sub>4</sub>) to 500 mL of DI water. This stock solution was then diluted with DI water to achieve target concentrations of 0.1, 0.2, 0.4, 1.0, 2.0, and 3.0 mg/L. These increments were selected based on the EPA’s SMCL of 1.0 mg/L for copper in drinking water, as well as to align with the increments that the test kits are marketed to measure (Table 2).

All concentrations were measured with CuVer Copper Reagent Power Pillows from Hach following Method 8506, the US EPA Bicinchoninate Method. This method could detect copper of concentrations 0.04 – 5.0 mg/L. Therefore, any sample with a concentration >5.0 mg/L had to be diluted with DI water before they could be measured with the spec. Method 8506 involved starting Program 135 Copper, Bicin on the spec. The sample was prepared by adding 10 mL of the sample to the sample cell and adding the contents of one CuVer 1 Reagent Powder Pillows to the cell. The cell was then swirled to mix and a timer started for 2 minutes. During those 2 minutes the blank was prepared. The blank consisted of 10 mL of the sample in a sample cell

without the addition of a powder pillow. This blank was put into the spec and used to zero the program. After 3 minutes the sample cell was ready to be read. Method 8506 reduces Cu<sup>2+</sup> to Cu<sup>+</sup> and then the Bicinchoninate reacts with the Cu<sup>+</sup> to form a purple colored complex.

This procedure was repeated for the other two source waters after adjusting for any background concentration of copper that was already present in the water sample, as explained previously in the section on iron. The target concentrations and measured copper concentrations were recorded (Table 9). Any target concentration that is repeated in Table 9 indicates that part of that experiment was done at a different time due to availability of the test kits that were ordered. Any concentration >0 mg/L listed in the row where the target concentration was 0 mg/L represents a background concentration in that water source. Solutions were made fresh immediately before testing the kits.

**Table 9.** Target and measured concentrations of copper solutions

Target Concentration (mg/L)	Measured Concentration using CuVer Copper Reagent Powder Pillows (mg/L)		
	DI Water	Amherst tap water	Mill River water
0.0	0.0	0.10	0.0
0.1	0.09	-	0.09
0.2	0.19	0.23	0.12
0.2	0.27	0.18	0.15
0.2	-	0.30	0.14
0.4	0.37	0.54	0.31
0.4	-	0.48	0.37
0.5	0.39	0.63	0.42
1.0	0.98	1.15	0.90
1.0	0.08	1.12	0.77
1.0	-	1.01	0.83
2.0	1.83	2.10	1.51
2.0	-	2.08	1.92
3.0	2.75	3.09	2.70
10.0	9.20	8.72	8.72

**Manganese.** A 2 mg/L stock solution of manganese was made by adding 0.0062 g of Manganese Sulfate Monohydrate (MnSO<sub>4</sub> · H<sub>2</sub>O) to 500 mL of DI water. This stock solution was then diluted with DI water to achieve target concentrations of 0.05, 0.1, 0.4, and 1.0 mg/L. These increments were selected based on the EPA’s SMCL of 0.05 mg/L for manganese in drinking water, as well as to align with the increments that the test kits were marketed to measure (Table 6).

All concentrations were measured following Hach’s Manganese LR PAN Method, Method 8149 which is capable of detecting manganese of concentrations 0.006 – 0.70 mg/L. Therefore, any sample with a concentration >0.70 mg/L had to be diluted with DI water before it could be measured in the spec. Method 8149 involved starting Program 290 Manganese, LR PAN on the spec. The blank was prepared by filling the sample cell with 10 mL of DI water. The sample was prepared by adding 10 mL of the sample to the sample cell. One powder pillow of Ascorbic Acid was added to each sample cell. Both cells were inverted to dissolve the powder. 12 drops of Alkaline Cyanide Reagent Solution were added to each cell and then the cells were swirled to mix. 12 drops of PAN Indicator Solution were added to each cell and then the cells were swirled to mix. After 2 minutes the blank was used to zero the spec and the sample was prepared to be read. Method 8149 adds ascorbic acid to reduce all oxidized forms of manganese to Mn<sup>2+</sup>. The alkaline cyanide reagent masks any potential interferences. Lastly, the PAN indicator forms an orange colored complex with Mn<sup>2+</sup>.

This procedure was repeated for the other two source waters after adjusting for any background concentration of manganese that was already present in the water sample. Similar to previous descriptions of iron and copper, the target concentrations and measured concentrations were recorded (Table 10). Concentrations >0 mg/L listed in the row where the target concentration was 0 mg/L represents a background concentration in that water source. Solutions were made fresh immediately before testing the kits.

**Table 10.** Target and measured concentrations of manganese solutions

Target Concentration (mg/L)	Measured Concentration using PAN Method (mg/L)		
	DI Water	Amherst tap water	Mill River water
0.0	0.0	0.009	0.008
0.05	0.044	0.057	0.062
0.10	0.099	0.108	0.111
0.40	0.20	0.404	0.394
2.00	2.45	7.70	8.20

**Fluoride.** Two fluoride stock solutions of 5 mg/L and 100 mg/L were made by adding 0.0056g and 0.1105g of Sodium Fluoride (NaF) respectively to 500 mL of DI water. This stock solution was then diluted with DI water to achieve target concentrations of 0.50, 1.0, 2.0, 4.0, 10, 25, and 50 mg/L. These increments were selected based on the EPA’s SMCL of 2.0 mg/L for fluoride in

drinking water, as well as to align with the increments that the test kits are marketed to measure (Table 7).

All concentrations were measured following Hach’s USEPA SPADNS 2 Method, Method 10225. This method could detect fluoride of concentrations 0.02 – 2.00 mg/L. Therefore, water sample with concentrations >2.00 mg/L had to be diluted with DI water before it could be measured in the DR Spec. Method 10225 involved starting Program 190 Fluoride on the spec. The sample was prepared by adding 10 mL of the sample to a sample cell. The blank was prepared by adding 10 mL of DI water to a sample cell. 2 mL of SPADNS 2 reagent was added to each cell. The cells were swirled to mix. After 1 minute the blank was used to zero the spec and the sample was read. Method 10225 involves fluoride reacting with a red zirconium dye. The fluoride combines with parts of the zirconium to form a colorless complex that bleaches the red color in an amount proportional to the fluoride concentration.

This procedure was repeated for the other two source water sources. Solutions were made fresh immediately before testing the kits.

**Table 11.** Target and measured concentrations of fluoride solutions

Target Concentration (mg/L)	Measured Concentration using SPADNS 2 Method (mg/L)		
	DI Water	Amherst tap water	Mill River water
0.0	0.0	-	-
0.50	0.52	1.14	0.57
1	0.87	1.85	1.46
2	1.70	2.0	3.80
4	3.40	4.60	5.60
5	3.80	4.60	4.60
10	10.6	15.8	17
25	36.0	26.0	20
50	53.0	58.0	59
100	179.0	-	-

#### 2.4.3. Experimental Procedure

First, the concentration of each solution was measured with the spec using the appropriate method. Then, 250mL glass beakers labeled by the target concentration of that parameter were filled with the solution. pH and temperature were also recorded for each solution, starting with the solution of lowest concentration and ending with the solution of highest concentration for that round of experimentation. At this point, each solution has a value recorded for target

concentration, measured concentration, pH, and temperature. For the DI water matrix, only concentration and temperature were recorded.

The final step was to take measurements using the OTS test kits. The solutions remained in the 250mL glass beakers and the instructions were followed for each kit (described previously), starting with the combination kits. A method may have included transferring the solution from the beaker into a provided sample vial. If not, the sample was transferred from the glass beaker to a 100 mL or 250 mL plastic cup. Each test kit brand was used five times on each solution, following that kit's specific instructions. After waiting the specified wait time for that test, the concentration result from the OTS kit test strip was recorded.

#### 2.4.4. Statistical Analysis

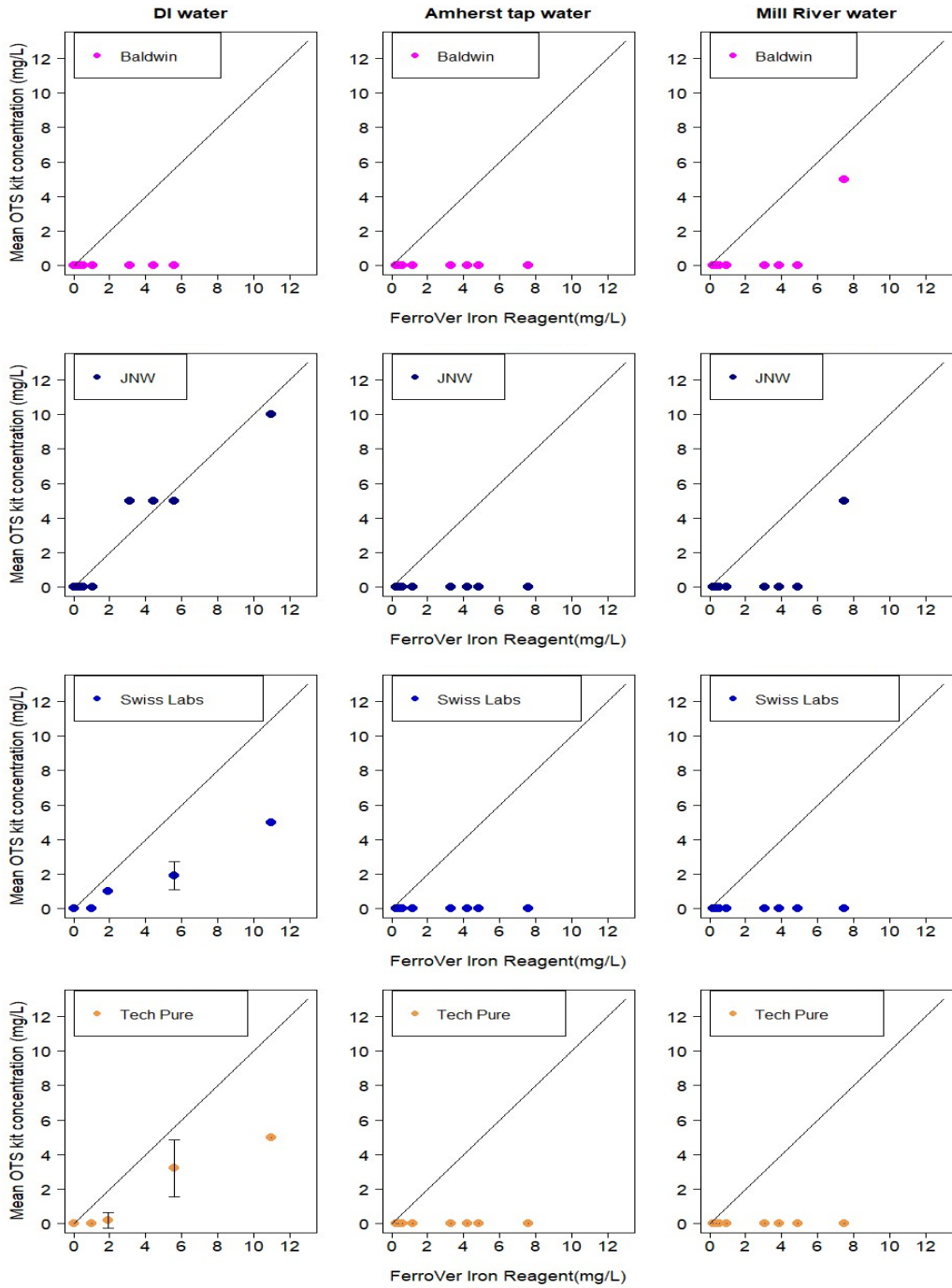
Results were entered into Microsoft Excel (2018). Calculations were performed and graphs were made using R (R Core Team 2018). More specifically, R was used to calculate the mean of the replicates at each concentration, standard deviation of the replicates at each concentration, and sensitivity of the test brands. Sensitivity was calculated as the number of true positives divided by the sum of the number of true positives and false negatives (Equation 1). To perform this analysis, quantitative results were transformed into binary values (positive or negative) based on whether the sample was equal to or greater than 0 mg/L, or whether the sample was above or below the respective SMCL value (this is explained further in the results section below).

$$Sensitivity = \frac{\# \text{ true positive}}{\# \text{ true positive} + \# \text{ false negative}} \times 100 \quad (\text{Equation 1})$$

## 2.5 Results

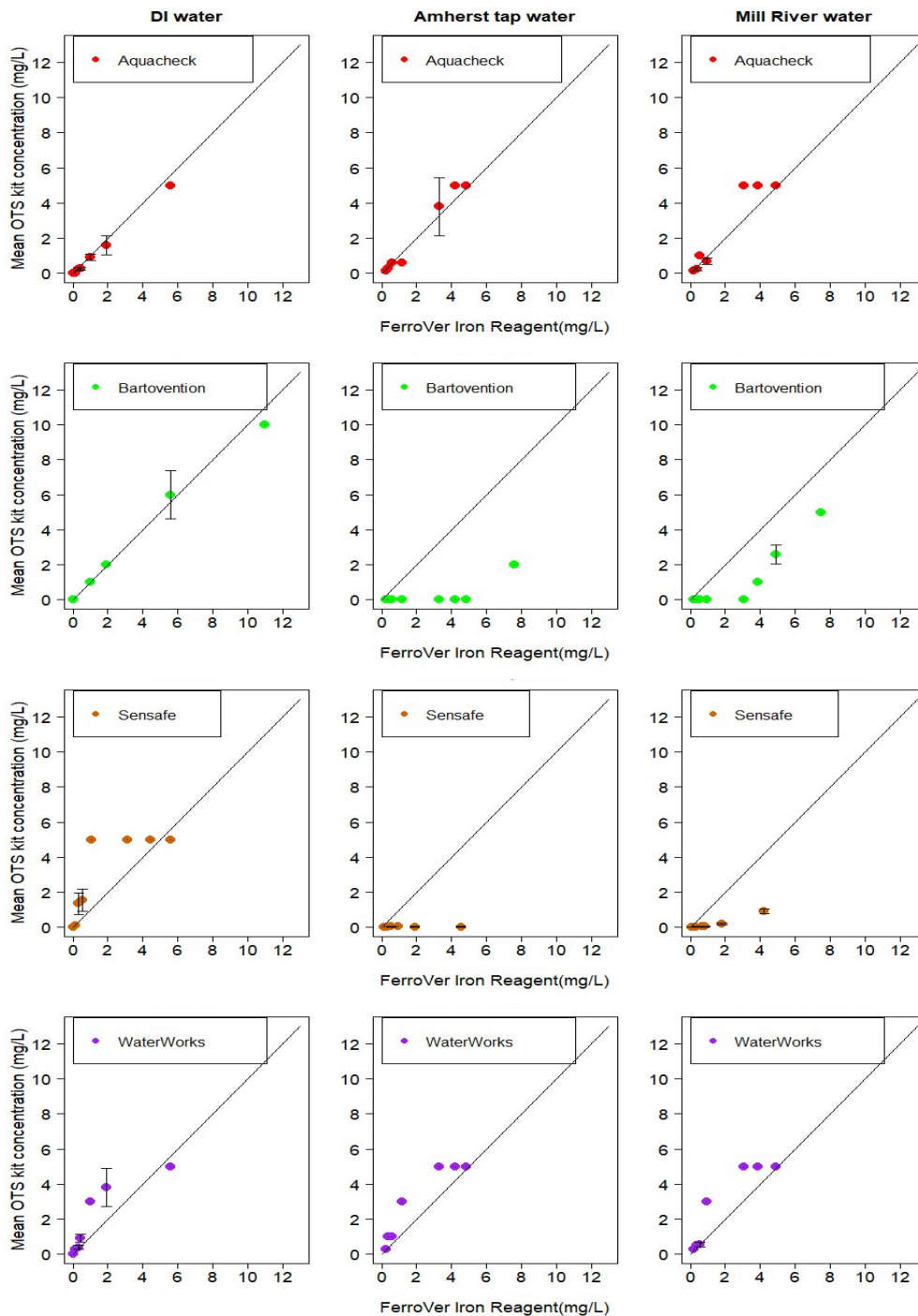
Results from all test kits at all concentrations are presented in Tables 14-25 located in the appendix. Most test kit-concentration measurements had a standard deviation of zero, meaning that the five replicate readings were the same. The control with DI water (nothing added), showed that test kits were able to measure 0 mg/L. However, many of the test kits at many concentrations were less than the actual concentration. Therefore, the results from the tests were typically precise (consistent between repeated tests) but not accurate. This was consistent in analyses performed in Amherst tap water and Mill River water as well.

### 2.5.1. Iron



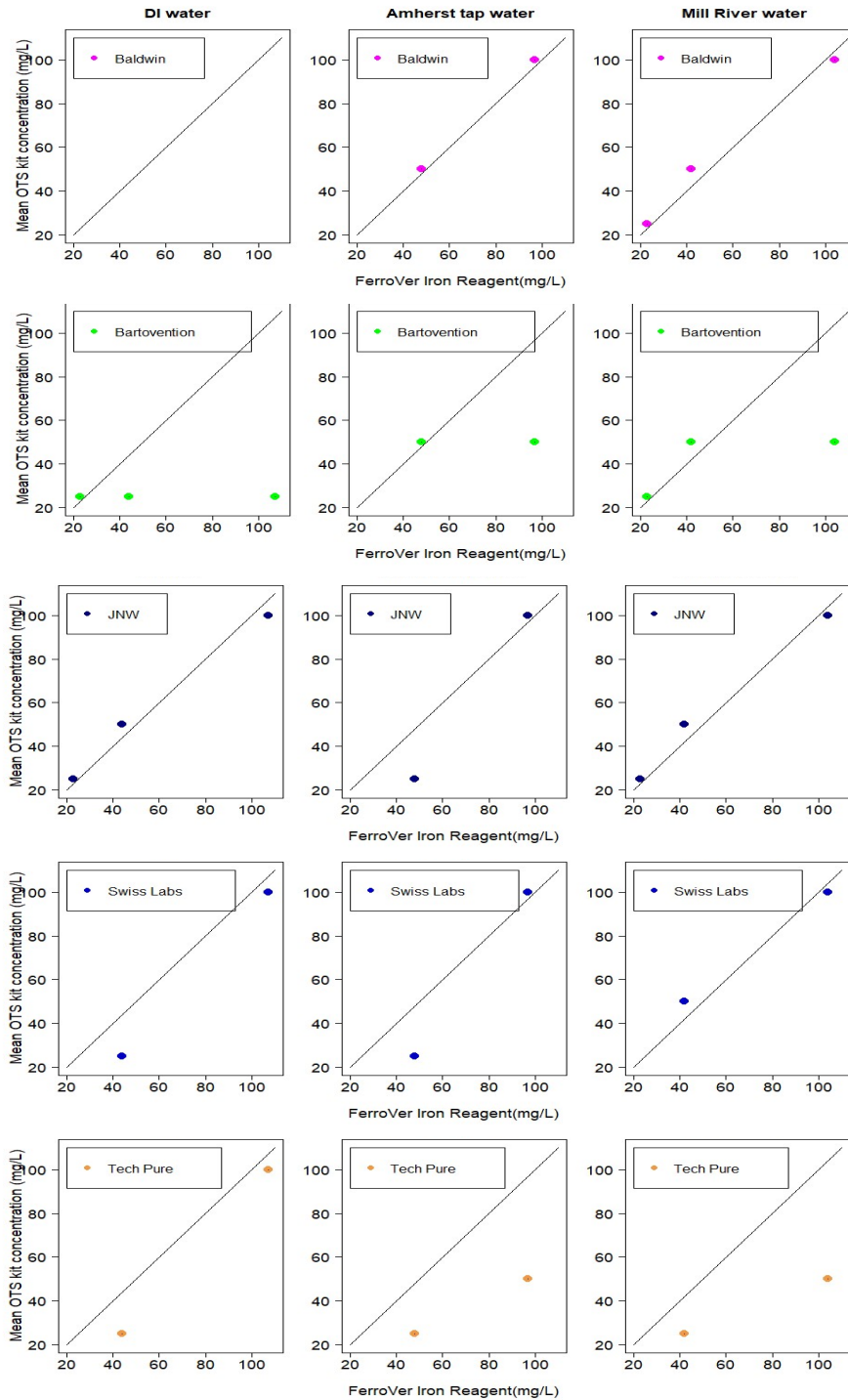
**Figure 2.** Iron (0-12 mg/L) as measured by combination test kit compared to the laboratory-obtained concentrations (FerroVer Iron Reagent). Each row represents a different kit and each column a different source water. Diagonal line represents a perfect match between FerroVer Iron Reagent Power Pillow measurements (taken as the laboratory-obtained accurate concentration) and the test kit. Each data point represents the mean of 5 trials per solution and 95% confidence interval; if no error bars are visible, all 5 tests yielded same results.

All combination test kits were able to accurately measure 0 mg/L in the DI water. Many of the combination test kits were able to detect iron in the DI water matrix, but not in Amherst tap water or Mill River water (Figure 2). The Baldwin Meadows kit was unable to detect iron in any of the source waters with the exception of a trial in Mill River water when the FerroVer Iron Reagent Powder Pillow measured 7.5 mg/L (although it underestimated this concentration by 5 mg/L). The JNW kit was unable to detect iron in the DI water matrix until the FerroVer Iron Reagent Powder Pillow measured 3.10 mg/L; at that point, the measured concentration were an overestimate at 5 mg/L. The remaining points were close to the perfect match line but still over or underestimated. However, the JNW kit did not detection any iron in the Amherst tap water and underestimated 7.5 mg/L as 5 mg/L in the Mill River water. Neither the Swiss Labs kit and the Tech Pure kit detected iron of any concentration in Amherst Tap water or Mill River water, although both were able to detect iron in DI water but underestimated all concentrations. The Swiss Labs and the Tech Pure kits were the only kits that had some measurements with a standard deviation not equal to zero.



**Figure 3.** Iron (0-12 mg/L) as measured by the iron only test kit compared to the laboratory-obtained concentrations (FerroVer Iron Reagent). Each row represents a different kit and each column a different source water. Diagonal line represents a perfect match between FerroVer Iron Reagent Power Pillow measurements (taken as the laboratory-obtained accurate concentration) and the test kit. Each data point represents the mean of 5 trials per solution and 95% confidence interval; if no error bars are visible, all 5 tests yielded same results.

The kits that measured only iron were all able to accurately measure 0 mg/L in the DI water source (Figure 3). The Aquacheck kit measured concentrations in the DI water matrix that were close to the laboratory-obtained concentrations (points close to the diagonal line). In the solutions made with Amherst tap water, the Aquacheck slightly underestimated at lower concentrations and overestimated at higher concentrations. In Mill River water, the Aquacheck kit slightly underestimated concentrations 0.20, 0.36, and 0.94 mg/L with results of 0.15, 0.24, and 0.68 mg/L, respectively, and overestimated 0.53 mg/L with the result 1.0 mg/L and, in all three measured concentrations >2 mg/L, measured 5 mg/L; however, considering that the kits could only measure discrete intervals, these values are close. The Bartovention kit measured concentrations nearly matching the laboratory-obtained concentration when in the DI water matrix: at concentrations 1.01, 1.91, and 5.58 mg/L, this kit gave the results 1, 2, and 6 mg/L respectively. However, in Amherst tap water, the Bartovention kit did not detect any iron until a concentration of 7.6 mg/L, which it estimated to be 2 mg/L. In Mill River water, the Bartovention kit underestimated all concentrations and measured 0 mg/L until 3.86 mg/L, which it estimated to be 1 mg/L. The Sensafe test kit had primarily overestimates in DI water and underestimations in Amherst Tap water and Mill River water. In DI water for concentrations of 0.33, 0.55, and 1.05 mg/L, the Sensafe kit overestimated the results to be 1.35, 1.55 and 5 mg/L respectively. In Amherst tap water, for concentrations of 0.1 and 0.33 mg/L the Sensafe kit underestimated both concentrations to be 0.02 mg/L. The trend of underestimation continued for the Sensafe kit in Mill River water. The Water Works kit often overestimated concentrations in all source waters.

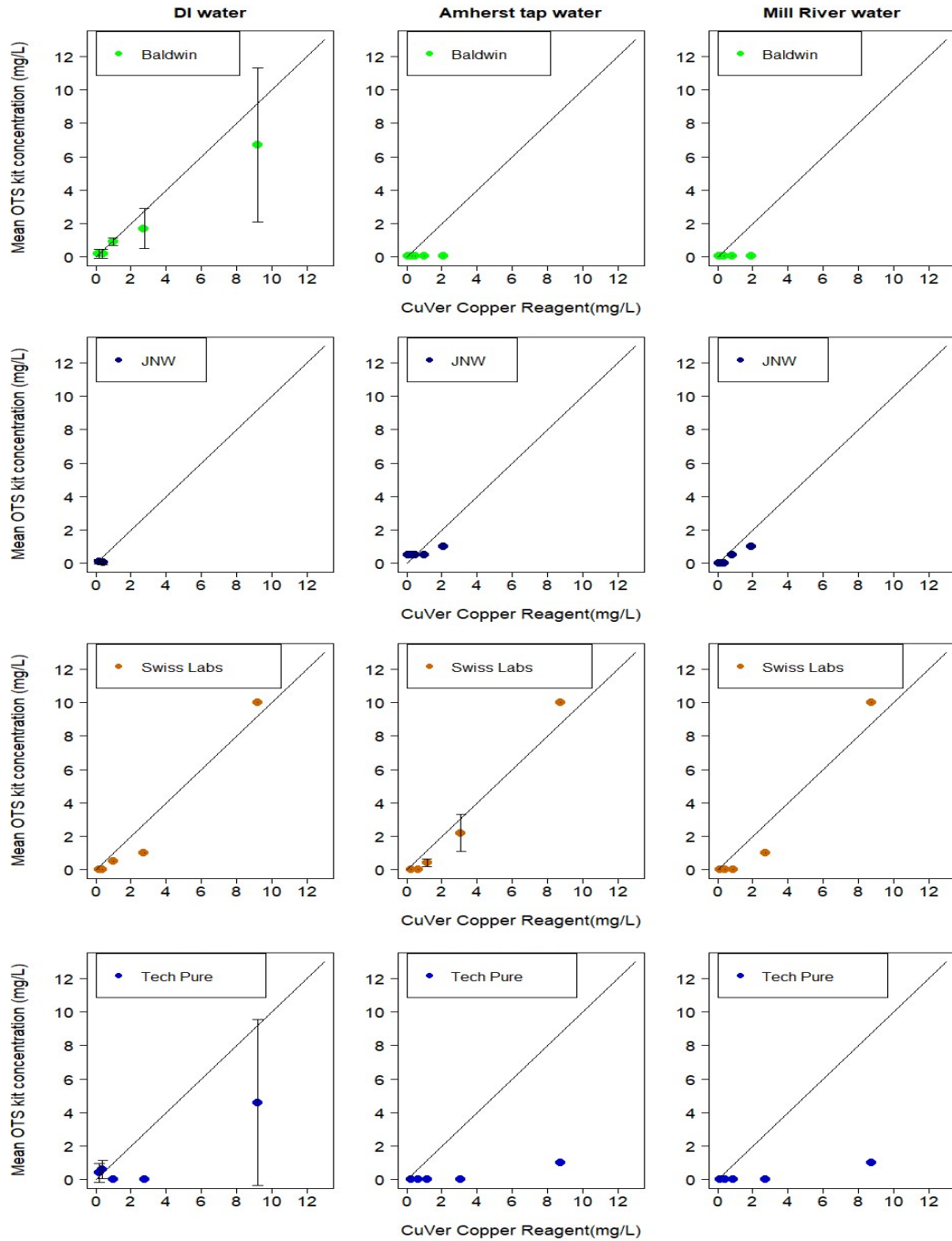


**Figure 4.** High range iron (20-100 mg/L) as measured by the iron only test kit and combination test kit compared to the laboratory-obtained concentration (FerroVer Iron Reagent). Each row represents a different kit and each column a different source water. Diagonal line represents a perfect match between FerroVer Iron Reagent Power Pillow measurements (taken as the laboratory-obtained accurate concentration) and the test kit. Each data point represents the mean of 5 trials per solution and 95% confidence interval; if no error bars are visible, all 5 tests yielded same results.

While Figures 2 and 3 illustrated the results from the combination kits and the iron only kits of concentration 0 – 12 mg/L of iron, many of the kits were advertised to measure much higher concentrations of iron. Therefore, Figure 4 shows the results from the five test kits (Baldwin Meadows, Bartovention, JNW, Swiss Labs, and Tech Pure) with a higher detection range of 20 – 100 mg/L. In Amherst tap water and Mill River water, the Baldwin Meadows kit was able to measure concentrations reasonably close to the laboratory-obtained results: for example, at 22.5 mg/L, in Amherst tap water it underestimated the concentration as 10 mg/L, while at 48 mg/L and 97 mg/L, the kit measured close values of 50 mg/L and 100 mg/L, respectively. In Mill River water at 23 mg/L and 42 mg/L, the kit slightly overestimated the results to be 25 mg/L and 50 mg/L respectively, and at 104 mg/L, estimated 100 mg/L. This kit performed better at concentrations 20 – 100 mg/L than it did at concentrations 0 – 12 mg/L for iron in water (Figure 2 and 3). The Bartovention kit is the only iron only kit with a detection range of 20 – 100 mg/L. In DI water, it measured 23, 44, and 107 mg/L all as 25 mg/L. In Amherst tap water, the kit underestimated many of the higher concentrations (therefore, the data points are not visible in Figure 4). In Amherst tap water, the Bartovention estimated 48 mg/L and 97 mg/L both as 50 mg/L and, similarly, in Mill River water, estimated 23 mg/L as 25 mg/L and 42 mg/L and 104 mg/L both as 50 mg/L. In DI water, the JNW kit provided very close estimates of 23, 44, and 107 mg/L as 25, 50, and 100 mg/L respectively. In Amherst tap water, the JNW kit underestimated 22.5 mg/L as 10 mg/L and 48 mg/L as 25 mg/L, but correctly estimated 97 mg/L as 100 mg/L. In Mill River water, the JNW kit also performed well, estimating 23, 42, and 104 mg/L to be 25, 50, and 100 mg/L, respectively. In DI water, Swiss Labs was similarly close to actual, with concentrations of 23, 44, and 107 mg/L measured by the kit as 10, 25, and 100 mg/L. However, in Amherst tap water, Swiss Lab greatly underestimated concentrations of 22.5 mg/L and 48 mg/L as 5 mg/L and 25 mg/L, respectively, but was nearly correct with 97 mg/L measuring as 100 mg/L. The Swiss Labs kit performed more variably in Mill River water, estimating 23, 42, and 104 mg/L to be 10, 50, and 100 mg/L respectively. The Tech Pure kit underestimated many concentrations, and therefore the data points do not appear on the figure: while in DI water it measured 23, 44, and 107 mg/L as 10, 25, and 100 mg/L, respectively, in Amherst tap water it estimated 22.5, 48, and 97 mg/L to be 8, 25, and 50 mg/L and, in Mill River water, estimated 23, 42, and 104 mg/L to be 10, 25, and 50 mg/L, respectively. Overall, the

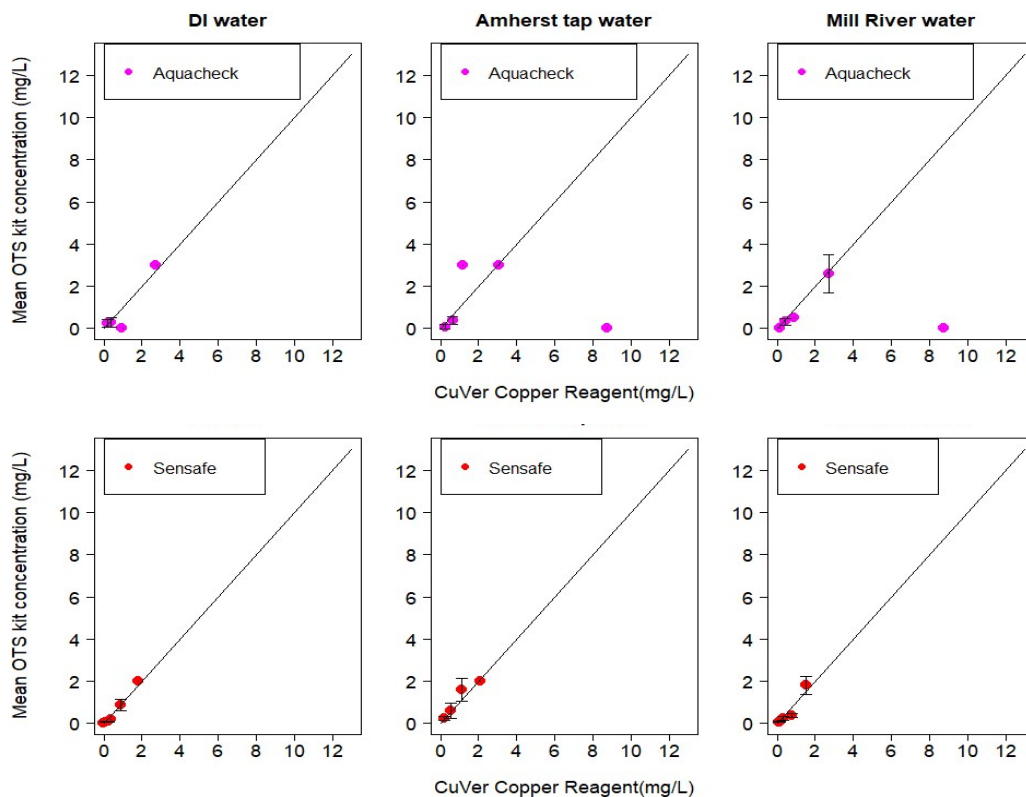
combination kits did not perform well at lower concentrations (Figure 2) but performed better at higher concentrations (Figure 4).

### 2.5.2. Copper



**Figure 5.** Copper concentrations (0-12 mg/L) as measured by combination test kit compared to the laboratory-obtained concentration (CuVer Iron Reagent). Each row represents a different kit and each column represents a different source water. Diagonal line represents a perfect match between CuVer Iron Reagent Power Pillow measurements (taken as the laboratory-obtained accurate concentration) and the test kit. Each data point represents the mean of 5 trials per solution and 95% confidence interval; if no error bars are visible, all 5 tests yielded same results.

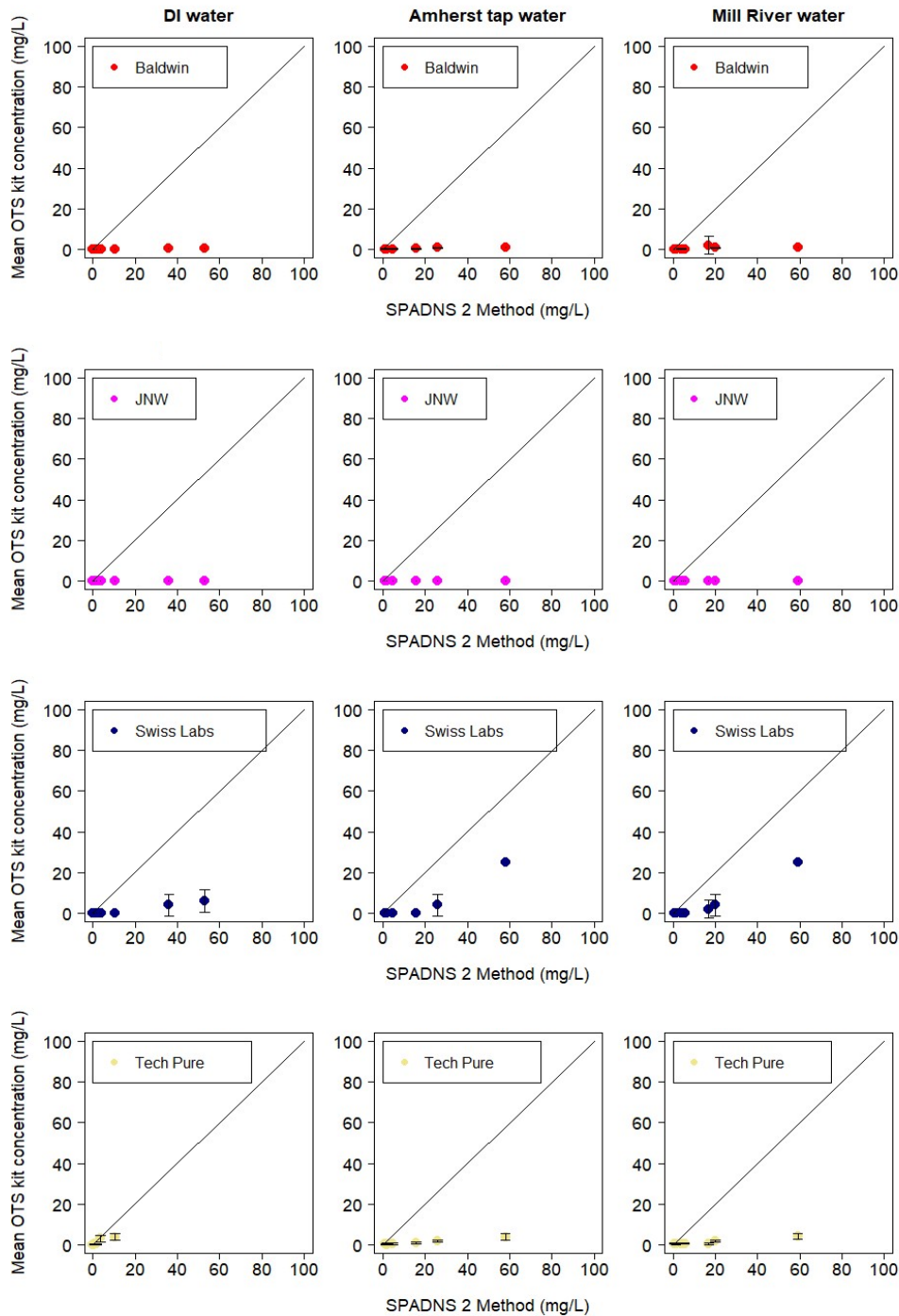
Four combination kits were tested on their ability to measure copper solutions ranging from 0-12 mg/L (Figure 5). The Baldwin Meadows kit was able to detect copper in DI water, however, in Amherst tap and Mill River water, it consistently underestimated all concentrations. For example, in DI water, the Baldwin Meadows kit measured 0.19, 0.39, 0.98, 2.75, and 9.20 mg/L as 0.2, 0.2, 0.9, 1.7, and 6.7 mg/L, respectively. However, in Amherst tap water and Mill River water, the Baldwin Meadows kit measured all concentrations to be 0.05 mg/L. Similarly, the JNW kit measured 0.19 mg/L and 0.39 mg/L as 0.1 mg/L and 0.05 mg/L, respectively, in DI water, but in Amherst tap water it measured all concentrations as 0.05 mg/L until 2.08 mg/L, which it estimated as 1 mg/L. In Mill River water, the JNW kit similarly measured all concentrations to be 0 mg/L until 0.83 mg/L and 1.92 mg/L, which it measured as 0.5 mg/L and 1 mg/L, respectively. The Swiss Labs kit provided somewhat better measurements at higher concentrations than lower concentrations in all water types. In DI water, it measured all concentrations to be 0 mg/L until 0.98, 2.75, and 9.2 mg/L, which it measured as 0.5, 1, and 10 mg/L, respectively. In Amherst tap water and Mill River water, the Swiss Labs kit similarly underestimated the lower concentrations and overestimated the higher concentrations. The Tech Pure kit in DI water estimated 0.19, 0.39, 0.98, 2.75, and 9.20 mg/L to be 0.4, 0.6, 0, 0, and 4.6 mg/L respectively, and performed similarly to the Swiss Labs kit in Amherst tap water and Mill River water by estimated all concentrations to be 0 mg/L until 8.72 mg/L, which it underestimated to be 1 mg/L.



**Figure 6.** Copper concentrations (0-12 mg/L) as measured by copper only test kit compared to the laboratory-obtained concentration (CuVer Iron Reagent). Each row represents a different kit and each column represents a different source water. Diagonal line represents a perfect match between CuVer Iron Reagent Power Pillow measurements (taken as the laboratory-obtained accurate concentration) and the test kit. Each data point represents the mean of 5 trials per solution and 95% confidence interval; if no error bars are visible, all 5 tests yielded same results.

Two copper only test kits were tested on copper concentrations from 0-12 mg/L (Figure 6). The Sensafe copper check performed better than the Aquacheck at low concentrations. In DI water, the Aquacheck estimated 0.19, 0.39, 0.98, and 2.75 mg/L to be 0.22, 0.28, 0, and 3 mg/L, respectively, yielding a mix of over and underestimates. In Amherst tap water, the Aquacheck kit produced similarly variable results (by estimating 0.23, 0.63, 1.15, 3.09, and 8.72 mg/L as 0.08, 0.38, 3, 3, and 0 mg/L, respectively) and, in Mill River water, it did not detect copper until 0.42 mg/L, which it underestimated slightly as 0.32 mg/L. Overall the Sensafe kit had a data point close to the perfect match line across all three water sources. The error bar on the Sensafe data point crosses the perfect match line, indicating at those points an over estimation or underestimation is possible.

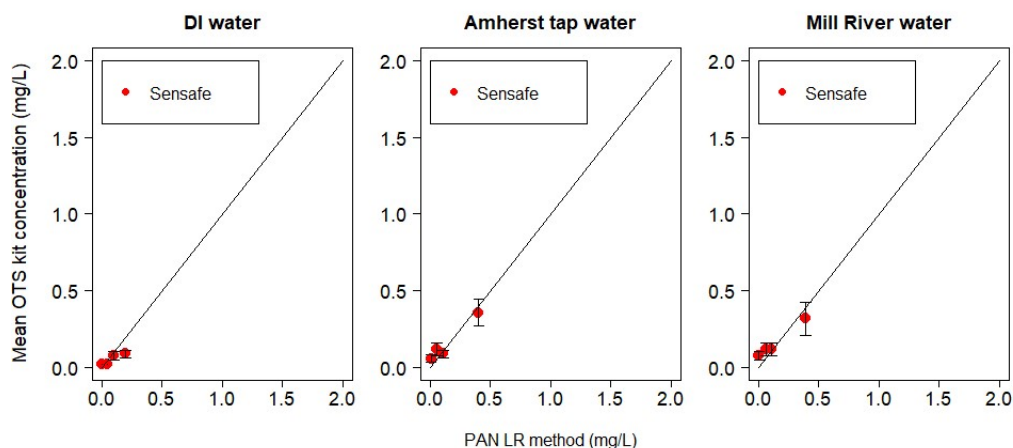
### 2.5.3. Fluoride



**Figure 7.** Fluoride concentrations (0-12 mg/L) as measured by combination test kit compared to the laboratory-obtained concentration (SPADNS 2 Method). Each row represents a different kit and each column a different source water. Diagonal line represents a perfect match between SPADNS 2 Method measurements (taken as the laboratory-obtained accurate concentration) and the test kit. Each data point represents the mean of 5 trials per solution and 95% confidence interval; if no error bars are visible, all 5 tests yielded same results.

Four combination test kits were tested for their ability to measure fluoride concentrations 0-12 mg/L (Figure 7). The only kit that was able to detect fluoride was the Tech Pure kit; the Baldwin Meadows, JNW, and Swiss Labs kits estimate all tested concentrations >0 mg/L as 0 mg/L. The Tech Pure kit underestimated all concentrations, but was able to detect fluoride >0 mg/L.

#### 2.5.4. Manganese



**Figure 8.** Manganese concentrations (0-2 mg/L) as measured by the manganese only test kit compared to the laboratory-obtained concentration (PAN LR Method). Each row represents a different kit and each column represents a different source water. Diagonal line represents a perfect match between PAN LR Method measurements (taken as the laboratory-obtained accurate concentration) and the test kit. Each data point represents the mean of 5 trials per solution and 95% confidence interval; if no error bars are visible, all 5 tests yielded same results.

One test kit, Sensafe, was analyzed for its ability to measure manganese solutions with concentrations 0-2 mg/L (Figure 8). This kit was able to estimate close to accurate concentrations in all three water sources.

#### 2.6. Sensitivity Analysis

Sensitivity was calculated by test kit brand for each source water to assess the ability of the kit to provide accurate results. To perform this analysis, we transformed the quantitative data into a binary measure: at greater than or equal to 0 mg/L (Figure 9) (i.e. whether, if there was any contaminant present in the water, the test kits were able to measure anything >0 mg/L), and at greater than or less than the EPA SMCL (Figure 10). While this does not provide information

about the error associated with test kits' results relative to laboratory-obtained concentration, this measure is designed to provide information about whether the test kits can provide useful yes/no information for users: how often they can provide meaningful data on detection and safety (as defined by the SMCL).

	<b>Spec = 0</b>	<b>Spec &gt; 0</b>
<b>Strip = 0</b>	<b>True negative</b>	<b>False negative</b>
<b>Strip &gt; 0</b>	<b>False positive</b>	<b>True positive</b>

**Figure 9.** Definitions for sensitivity calculation based on detection

	<b>Spec &lt; SMCL</b>	<b>Spec ≥ SMCL</b>
<b>Strip &lt; SMCL</b>	<b>True negative</b>	<b>False negative</b>
<b>Strip ≥ SMCL</b>	<b>False positive</b>	<b>True positive</b>

**Figure 10.** Definitions for sensitivity calculation based on the USEPA's SMCLs

$$Sensitivity = \frac{\# \text{ true positive}}{\# \text{ true positive} + \# \text{ false negative}} \times 100 \quad (\text{Equation 1})$$

Under the first definition, a true positive was defined as a test strip reporting any detection (Figure 9). Therefore, if a test strip reported any concentration >0 mg/L when the solution was not a DI control (therefore, the actual concentration was >0 mg/L), that measurement was categorized as a true positive. A false negative was defined as a test strip reporting a concentration of 0 mg/L when the actual concentration was >0 mg/L. The sensitivity of each test kit brand to detection (>0 mg/L) was then calculated using Equation 1.

Under the second definition, a true positive was defined as the strip reporting a concentration greater than or equal to that constituent's SMCL and the actual concentration being greater than or equal to that constituent's SMCL (i.e., they both had the same binary result)

(Figure 10). Therefore, a false negative was when the test strip reported a concentration less than the SMCL of a constituent, but the actual concentration was greater than the SMCL. The sensitivity of each test kit brand to detect the SMCL was then calculated using Equation 1.

2.6.1. Sensitivity of test kits measuring iron

	DI Water	Amherst Tap Water	Mill River Water	Total
Water Works	0.86	1.00	1.00	0.95
Aquacheck	0.71	1.00	1.00	0.90
Sensafe	1.00	0.86	0.83	0.90
Bartovention	1.00	0.57	0.86	0.81
JNW	1.00	0.43	0.57	0.67
Swiss Labs	0.86	0.43	0.43	0.57
Tech Pure	0.74	0.43	0.43	0.53
Baldwin	0.00	0.43	0.57	0.33

**Figure 11.** Sensitivity analysis of kits measuring iron based on ability to detect. Each row represents a different test kit and each column represents a different water matrix. The total column takes the average of the three water matrix values.

	DI Water	Amherst Tap Water	Mill River Water	Total
Bartovention	1.00	0.57	0.86	0.81
Water Works	0.71	0.86	0.86	0.81
Aquacheck	0.63	0.86	0.80	0.76
JNW	1.00	0.43	0.57	0.67
Swiss Labs	0.86	0.43	0.43	0.57
Tech Pure	0.74	0.43	0.43	0.53
Baldwin	0.00	0.43	0.57	0.33
Sensafe	0.86	0.00	0.14	0.33

**Figure 12.** Sensitivity analysis of kits measuring iron based on EPA’s SMCL of 0.3 mg/L of iron. Each row represents a different test kit and each column represents a different water matrix. The total column takes the average of the three water matrix values.

The sensitivity of iron measurements made by each kit in each water source was calculated based on their ability to detect anything (Figure 11) and their ability to detect the EPA’s SMCL of 0.3 mg/L of iron in drinking water respectively (Figure 12). The Water Works kit performed better than all other kits (Figure 11). Interestingly, the Water Works kit had a sensitivity of 0.86 in the DI water matrix and a sensitivity of 1.0 in the other two water sources. The Aquacheck kit was also more sensitive in the source waters versus the DI water (perfect

sensitivity in tap water and Mill River water, but 0.71 in DI water). The top four best performing kits were those that measured only iron, and the four worst performing kits were the combination kits (Figure 11).

In their ability to detect whether a solution was above or below the SMCL (Figure 12), the Bartovention kit performed much better than in the previous sensitivity analysis (Figure 11). The Sensafe kit severely underperformed by this measure, dropping from performing well in the detection sensitivity analysis to performing the worst in the SMCL performance analysis. However, the Water Works kit ranked first and second out of all the kits according to both measures in both tap water and river water.

### 2.6.2. Sensitivity of test kits measuring copper

	DI Water	Amherst Tap Water	Mill River Water	Total
<b>Sensafe</b>	1.00	0.80	1.00	0.93
<b>Baldwin</b>	0.76	1.00	1.00	0.92
<b>Aquacheck</b>	0.52	0.68	0.60	0.60
<b>Swiss Labs</b>	0.60	0.56	0.40	0.52
<b>JNW</b>	0.12	1.00	0.40	0.51
<b>Tech Pure</b>	0.40	0.20	0.20	0.27

**Figure 13.** Sensitivity analysis of kits measuring copper based on ability to detect. Each row represents a different test kit and each column represents a different water matrix. The total column takes the average of the three water matrix values.

	DI Water	Amherst Tap Water	Mill River Water	Total
<b>Swiss Labs</b>	0.40	0.40	0.40	0.40
<b>Aquacheck</b>	0.20	0.40	0.20	0.27
<b>Sensafe</b>	0.20	0.40	0.20	0.27
<b>Tech Pure</b>	0.20	0.20	0.20	0.20
<b>JNW</b>	0.00	0.20	0.20	0.13
<b>Baldwin</b>	0.32	0.00	0.00	0.11

**Figure 14.** Sensitivity analysis of kits measuring copper based on EPA’s SMCL of 1.0 mg/L of copper. Each row represents a different test kit and each column represents a different water matrix. The total column takes the average of the three water matrix values.

Figures 13 and 14 show the results of the sensitivity analysis on the test kits that were used to test the copper solutions based on their ability to detect any copper (>0 mg/L), and their

ability to detect the EPA’s SMCL of 1.0 mg/L respectively. Comparing the two figures, the same trend as observed in the iron results occurs in which the Sensafe kit performs well in the sensitivity analysis of detection, but then performs worse in the sensitivity analysis of the SMCL. In both analyses, some of the combination kits outperformed the kits that measure only copper concentration. The Baldwin Meadows kit outperformed the Aquacheck kit in the detection sensitivity analysis, but then performed the worst out of all the kits in the SMCL sensitivity analysis. Notably, the highest sensitivity value out of any of the kits based on the SMCL sensitivity analysis was only 0.40, which is very low; most kits seemed to be designed well for detection, but not at the level of the SMCL.

2.6.3. Sensitivity of test kits measuring fluoride

	DI Water	Amherst Tap Water	Mill River Water	Total
Tech Pure	0.56	0.63	0.80	0.66
Baldwin	0.33	0.33	0.27	0.31
Swiss Labs	0.22	0.16	0.18	0.19
JNW	0.11	0.00	0.00	0.04

**Figure 15.** Sensitivity analysis of kits measuring fluoride based on ability to detect. Each row represents a different test kit and each column represents a different water matrix. The total column takes the average of the three water matrix values.

	DI Water	Amherst Tap Water	Mill River Water	Total
Tech Pure	0.33	0.22	0.20	0.25
Swiss Labs	0.22	0.16	0.18	0.19
Baldwin	0.11	0.00	0.02	0.04
JNW	0.11	0.00	0.00	0.04

**Figure 16.** Sensitivity analysis of kits measuring fluoride based on EPA’s SMCL of 2.0 mg/L of fluoride. Each row represents a different test kit and each column represents a different water matrix. The total column takes the average of the three water matrix values.

Figures 15 and 16 show the results of the sensitivity analysis on the test kits that were used to test the fluoride solutions based on their ability to detect anything and their ability to detect the EPA’s SMCL of 2.0 mg/L of fluoride in drinking water respectively. Overall, the kits performed better in the detection sensitivity analysis than in the SMCL sensitivity analysis. The

Technically Pure kit performed the best in both analyses and the JNW kit performed the worst in both analyses. Additionally, the kits typically performed better in the DI water matrix than in the other two waters.

#### 2.6.4. Sensitivity of test kits measuring manganese

	DI Water	Amherst Tap Water	Mill River Water	Total
Sensafe	0.80	1.00	1.00	0.93

**Figure 17.** Sensitivity analysis of kit measuring manganese based on ability to detect. Each row represents a different test kit and each column represents a different water matrix. The total column takes the average of the three water matrix values.

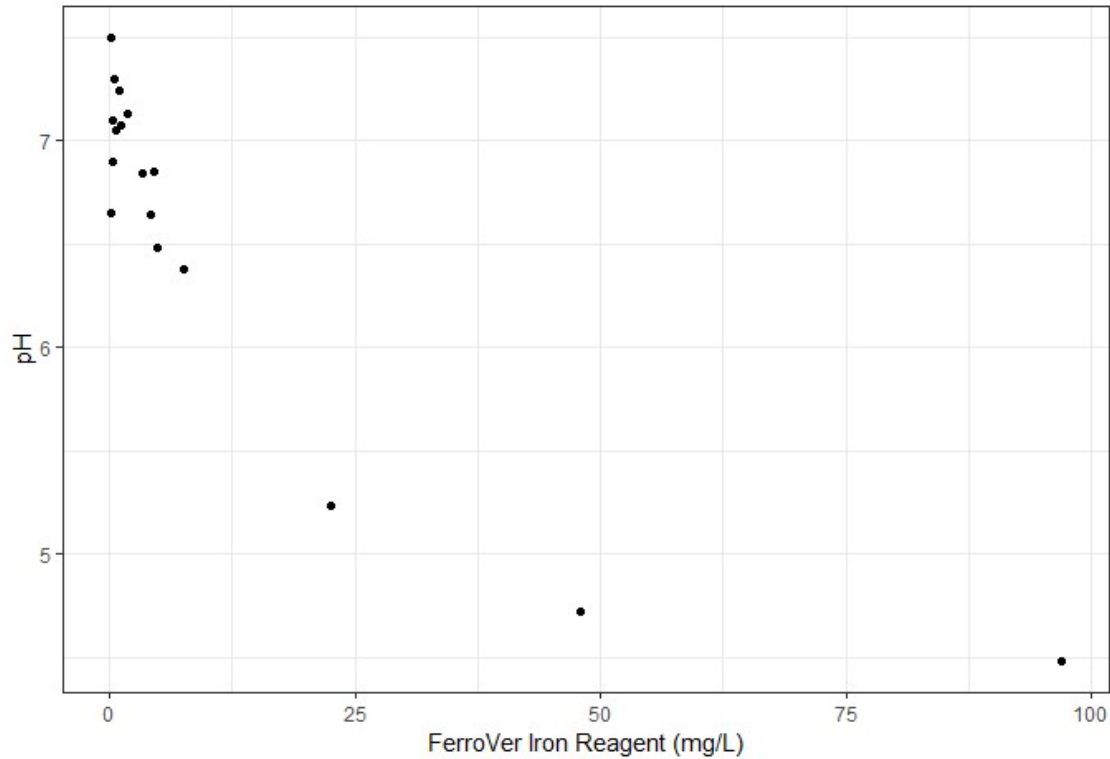
	DI Water	Amherst Tap Water	Mill River Water	Total
Sensafe	0.60	0.80	0.80	0.73

**Figure 18.** Sensitivity analysis of kits measuring manganese based on EPA’s SMCL of 0.05 mg/L of manganese. Each row represents a different test kit and each column represents a different water matrix. The total column takes the average of the three water matrix values.

Figures 17 and 18 show the results of the sensitivity analysis on the test kit that was used to test the manganese solutions based on its ability to detect anything and its ability to detect the EPA’s SMCL of 0.05 mg/L of manganese in drinking water respectively. The Sensafe kit had a higher sensitivity in the analysis based on detection than in the analysis based on the SMCL. It is worth noting that in the analysis based on detection the Sensafe kit performed better in Amherst tap water and Mill River water than it did in DI water. While there were no other kits on the market available for comparison, these sensitivity values (0.93 and 0.73) are quite high.

#### 2.6.5. Influence of pH on results

As the concentration of iron increased, the pH decreased (Figure 19). pH will affect the solubility of iron, therefore could affect the form of iron that is in the water, which would matter depending on the type of iron the kit can measure. Further research would need to be conducted in order to determine how pH affect the test kits and their methods.



**Figure 19.** Iron concentrations (0-100 mg/L) in the Amherst tap water matrix measured by the FerroVer Iron Reagent compared to the recorded pH.

## 2.7. Ranking

We ranked the test kits by parameter based on the laboratory results and the sensitivity analysis presented in Table 12. Ultimately, the sensitivity analysis based on a kit’s ability to detect the SMCL was deemed the most important and most useful for a citizen scientist and was therefore used to produce this ranking.

Notably, no one brand or kit performed consistently well across multiple water quality parameter; the same brand in some cases performed well for one parameter (e.g. Sensafe’s manganese kit) and poorly for another (e.g. Sensafe’s iron kit), or the same combination kit was able to adequately measure one parameter (e.g. Baldwin’s copper results) but not another (e.g. Baldwin’s iron results). This ranking was hopefully going to show some trends that indicated a few kits that were better overall than the others. Unfortunately, that trend did not emerge and there appears to be no one brand that could be turned to all the time.

**Table 12.** Final Ranking of test kit brands based on parameter

<b>Iron</b>	<b>Copper</b>	<b>Fluoride<sup>a</sup></b>	<b>Manganese</b>
1. Water Works	1. Sensafe	1. Tech Pure	1. Sensafe
2. Aquacheck	2. Baldwin	2. Swiss Labs	
3. Bartovention	3. Swiss Labs	3. Baldwin	
4. JNW	4. Aquacheck	4. JNW	
5. Sensafe	5. JNW		
6. Swiss Labs	6. Tech Pure		
7. Tech Pure			
8. Baldwin			

a – The highest sensitivity was 0.66 for detection 0.25 for the SMCL definition; these are notably low.

## 2.8. Discussion

*Precision vs. accuracy.* All measurements with the test kits were performed five times (with five different strips in the same solution). The vast majority of repeated measurements obtained from each of the five tests were the same. Therefore, the results from the test kits were generally highly precise; however, the concern was accuracy in terms of their ability to align with laboratory-obtained values.

*Water matrices.* Kits measuring iron were typically able to better approximate the laboratory-obtained concentration in the DI water matrix as compared to in the Amherst tap water and Mill River water. In general, there were no noticeable differences in performance between tests performed in Amherst tap water compared to Mill River water. In DI water there would be no other constituents present to interfere with the test kit. In contrast, there could be other constituents present in the other two water sources that could interfere with a test kits' ability to detect a certain parameter. Similarly, kits measuring copper followed the same trend of being better able to approximate laboratory-obtained concentration in DI water than in Amherst tap water and Mill River water. The evaluated kits measuring fluoride were unable to accurately measure fluoride concentrations (generally measuring no fluoride even when it was present) in all three water matrices; one was able to detect fluoride, but not accurately. In the kits that performed poorly across all three matrices, it is possible that the problem is with the kit's

methods and not the interference of other constituents, since DI water would not contain anything that would cause interference. The kit testing manganese performed better in Amherst tap water and Mill River water than in DI water; this method was a direct modification of the Hach PAN LR spec method, with the steps and reagents closely matching each other.

*Combination vs. parameter-only kits.* The test kits that measured only single parameters generally performed better than the combination kits when measuring iron solutions. However, it is important to consider the intervals at which each kit measures. Therefore, a concentration measured by a test kit that was very different from the laboratory-obtained concentration may have been limited not by the intensity of color it produced, but by the defined intervals given on the bottle. When examining the sensitivity results based on the SMCLs, some of the combination kits outperformed the parameter-only kits: for example, while the copper-only kits may be more accurate (have values closer to the laboratory-obtained results), the sensitivity analysis reveals that some of the combination kits outperform the copper only kits when assessing their ability to measure whether the concentration was greater or less than the SMCL. Only combination kits were used for the fluoride and manganese solutions, so it was not possible to compare these differences.

Notably, two of the test kits measuring only iron instructed users to add a reducing agent to the water sample before testing. This reducing agent addition would ensure that any ferrous iron is reduced to ferric iron. Therefore, the iron in the sample will be in the same form, and the test measures that iron. None of the combination kits required the addition of a reducing agent. Therefore, the combination kit would need to be measuring multiple forms of iron, or it missed measuring some of the iron in the sample. The kits that required this addition consistently performed well.

*Interpreting measurements.* The specificity analysis resulted in data that explains the probability that a test kit would produce a positive result. A true positive was then either that a test would report any concentration  $>0$  mg/L if there was any concentration present  $>0$  mg/L, or that a test would report a value above the SMCL if it was truly above the SMCL. For the purposes of citizen science in this context, the ability of a test kit to meaningfully provide information about a water sample's compliance with the SMCL is potentially more important than accuracy of a kit or the ability to detect anything above 0 mg/L. In the context of drinking water, often a citizen scientist's primary concern is the SMCL or MCL to know whether or not

their water is safe. However, accuracy may be a more important consideration if these data were to be used for scientific studies of water quality.

Overall, the combination kits simplify the science behind drinking water to a point that could be dangerous. Especially since they have consistently low sensitivities in relation to the SMCLs for the four parameters. That means that if a water sample is above the SMCL it is unlikely that will register with the combination kit. Therefore, I would recommend a single parameter kit to a citizen scientist over a combination kit. The single parameter kit allows for more details about that parameter to be provided to the consumer. Many of the single parameter kits model themselves after the laboratory methods, which would give a citizen scientist more insight into the true methods. The iron kits and manganese kit both required the addition of a powder pillow, just as it is done in the lab. However, more steps in the test kits does allow for more error on the part of the citizen scientist.

*Limitations.* For these experiments, tap and river water were sometimes collected on different days, which would result in different background constituents (important for potential interference), and concentrations of the parameter of interest. Although background concentrations of the measured parameters were considered at each time water was collected, it would be more consistent to collect all the water at one time. Additionally, when collecting tap water, there should be a consistent amount of time to flush the tap. Flushing the tap will get rid of any constituents that have been sitting and building up in the pipes overnight. Test strips were also read at different times of the day, therefore the lighting in the room would have been slightly different which may have slightly varied the results. There was also the possibility of reader error when interpreting the colors on the color chart.

*Future work.* This research provides a method for assessing whether and, if so, which commercially sold test kits would be most useful for a citizen scientist. However, there are more factors to consider in order to account for different drinking water conditions and usability by a citizen scientist (which is further addressed in Chapter 3). Some other experimentations and questions to be thought of include:

- Testing solutions at each concentration with varying pHs and temperatures to account for different drinking water conditions.
- Testing other sources of water that might be used by citizen scientists, such as groundwater.

- During experiments, vary some of the key method procedures to reproduce mistakes likely made by citizen scientists, including waiting more or less time than instructed to dip the strip in the water or matching the color, varying sample volume.
- Investigate what is in the source waters to see what may be causing interference with the test kits.
- Measure parameter of interest concentrations use other methods (e.g. a ICP-MS) to account for the possibility of errors with the Hach methods used.
- Evaluate the ability of combination test strips to measure more parameters than the four evaluated in this study, or to accurately measure parameters in parallel.

Additionally, after seeing most of the five replicates produce the same results (i.e., precise) it would be plausible to decrease the number of replicates to two and include a procedure if the two replicates are vastly different. There would also need to be some investigating into the chemistry of the strips themselves. The results from this work show significant differences in results from test kit to test kit across the same concentration. The chemistry on the strip is the most likely cause for this difference, therefore it would be useful to investigate it and determine the methods the strips are using.

## Chapter 3: Study Group Case Study

### 3.1. Overview

Following the laboratory studies, we asked student groups to use the test kits as representatives of potential citizen scientists. Undergraduate student groups from the College of Engineering at the University of Massachusetts Amherst were invited to participate in this portion of the research project. A data sheet was given to the students to record their results individually and in groups of two or more. A survey was also included on this data sheet to get input from these citizen scientists on their experience using the test kits.

### 3.2. Research Objectives

The objectives of the student group case study portion of this research project were to:

- Determine whether citizen scientists could successfully use these selected test kits.
- Learn citizen scientists' satisfaction, or dissatisfaction, with the test kits and their methods.
- Provide an educational opportunity for undergraduate engineering organizations.

### 3.3. Methods

#### *3.3.1. Study Design*

The study was determined to be exempt by the University of Massachusetts Amherst Institutional Review Board (IRB). A protocol was submitted for approval (Protocol ID 2019-5435). The approved protocol and consent documents are included in the Appendix.

For this portion of the study, only iron test kits were included: Sensafe, Aquacheck, and Swiss Labs. This ensured that citizen scientists were able to use a combination kit and a kit that measured only iron concentration. The Aquacheck kit performed well in both sensitivity analyses, however it was one of the few kits that performed worse in the DI water matrix compared to the other water matrixes. The Sensafe kit performed well in the sensitivity analysis for detection but, in contrast, performed poorly in the sensitivity analysis for the SMCL. The Swiss Labs performed the opposite and performed well in the sensitivity analysis for the SMCL while underperforming in the sensitivity analysis for detection. The citizen science data provided more insight into these kits and gave more factors to consider for ranking.

Two water sources were used: DI water and Amherst tap water. During the laboratory portion the kits performed similarly in Amherst tap water and Mill River water, therefore, the Mill River water matrix was eliminated for the citizen scientist portion.

### 3.3.2. Group Selection and Data Collection

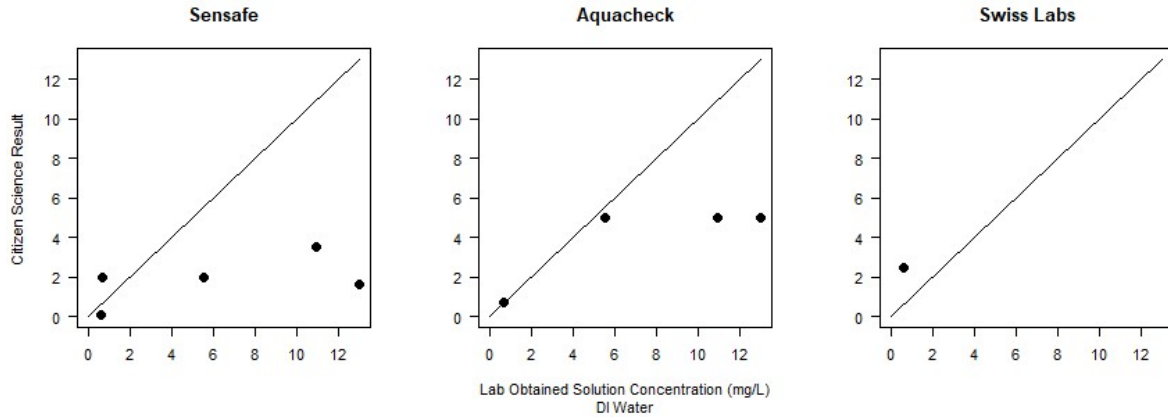
During meetings for the UMass chapters of the Society of Women Engineers (SWE) and Engineers Without Borders (EWB), students were asked to volunteer to participate in this study. A total of 58 students participated during various group meetings in April 2019. Prior to the meetings, solutions were made fresh. The laboratory-obtained values for the concentrations of these solution were recorded and then the solutions were put into color coded containers (Table 13).

**Table 13.** Color codes and lab-obtained solution concentrations of water samples brought to meetings with EWB and SWE for citizen science testing.

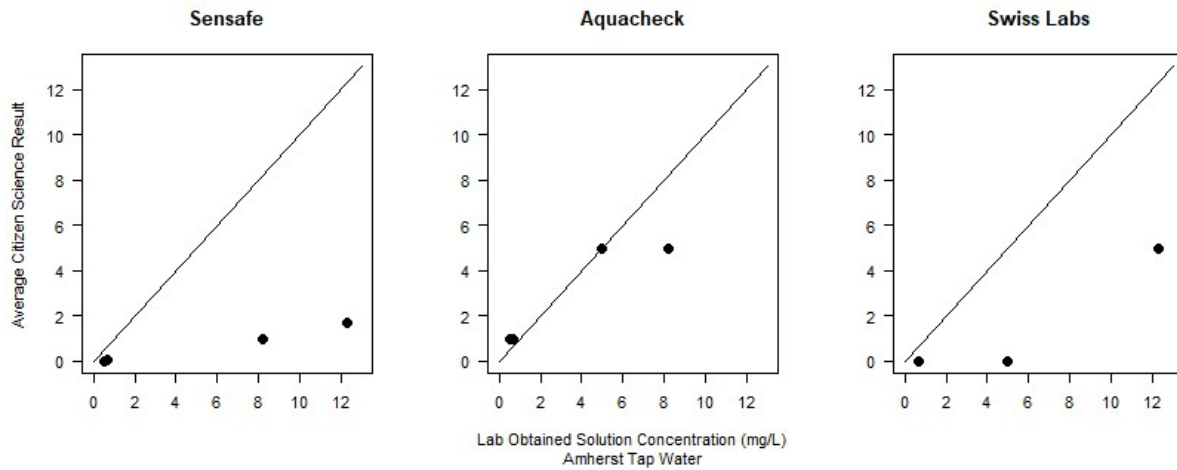
EWB meeting			
DI Water		Amherst Tap Water	
Color Code	Concentration	Color Code	Concentration
teal	0.56	orange	0.56
blue	5.00	yellow	13.00
red	10.60	purple	10.10
pink 2	118.00	green	5.50
		pink 1	96.00
SWE meeting			
DI Water		Amherst Tap Water	
Color Code	Concentration	Color Code	Concentration
yellow	10.90	red	12.30
purple	12.00	blue	8.20
orange	0.63	teal	0.70

### 3.4. Results

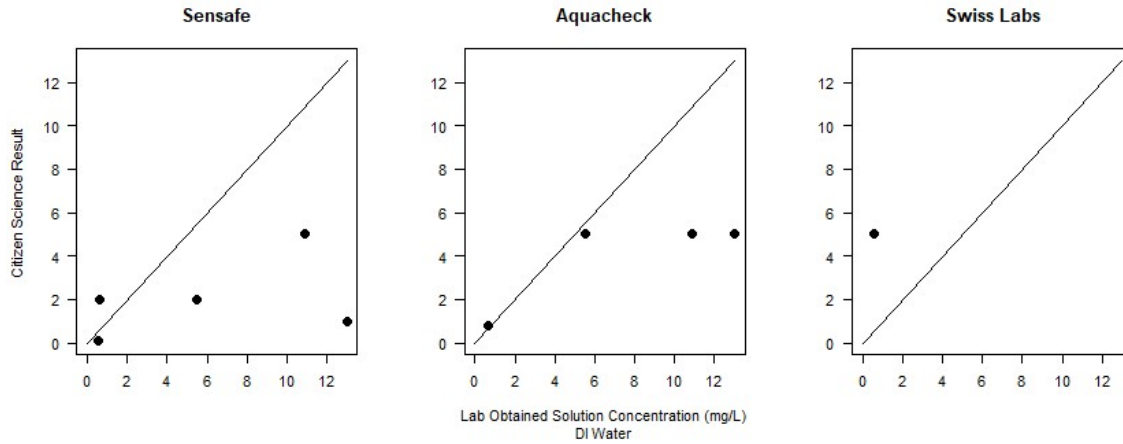
Although we provided the citizen scientists with solutions over 100 mg/L of iron, most of the data fell within the 0-12 mg/L range, so these results are presented here (also, solutions with concentrations >100 mg/L of iron were noticeably brown in color). We separated the data by water source, test kit, and individual or group results. To compare the citizen science data to the laboratory-obtained data, we took the average citizen science-obtained result per concentration, just as we took the average of five replicates in the laboratory portion of this study. We also analyzed the citizen scientists' responses to the questions regarding the test kits' clarity, ease of use, and results.



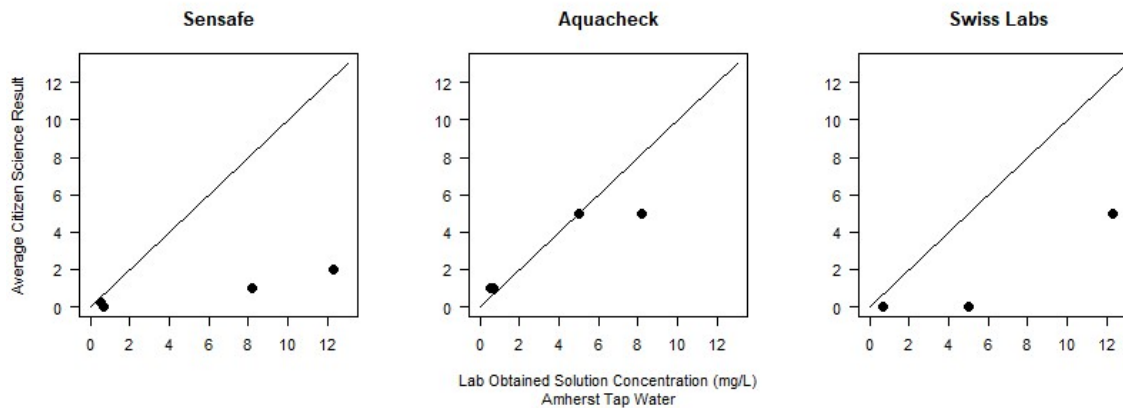
**Figure 20.** Iron concentrations (0-12 mg/L) as measured individually by the citizen scientists in the DI water matrix. Each column represents a different test kit. The average individual citizen scientist result on the y axis is compared to the laboratory-obtained value (FerroVer Iron Reagent) on the x axis. Diagonal line represents a perfect match between FerroVer Iron Reagent Power Pillow measurements (taken as the laboratory-obtained accurate value) and the citizen scientists' results.



**Figure 21.** Iron concentrations (0-12 mg/L) as measured individually by the citizen scientists in the Amherst tap water matrix. Each column represents a different test kit. The average individual citizen scientist result on the y axis is compared to the laboratory-obtained concentration (FerroVer Iron Reagent) on the x axis. Diagonal line represents a perfect match between FerroVer Iron Reagent Power Pillow measurements (taken as the laboratory-obtained accurate concentration) and the citizen scientists' results.



**Figure 22.** Iron concentrations (0 -12 mg/L) as measured in groups by the citizen scientists in the DI water matrix. Each column represents a different test kit. The average group citizen scientist result on the y axis is compared to the laboratory-obtained concentration (FerroVer Iron Reagent) on the x axis. Diagonal line represents a perfect match between FerroVer Iron Reagent Power Pillow measurements (taken as the laboratory-obtained accurate concentration) and the citizen scientists' results.



**Figure 23.** Iron concentrations (0 -12 mg/L) as measured in groups by the citizen scientists in the Amherst tap water matrix. Each column represents a different kit. The average group citizen scientist result on the y axis is compared to the laboratory-obtained concentration (FerroVer Iron Reagent) on the x axis. Diagonal line represents a perfect match between FerroVer Iron Reagent Power Pillow measurements (taken as the laboratory-obtained accurate concentration) and the citizen scientists' results.

The individual and group results yielded very similar results. However, there were differences when comparing the citizen science measured concentrations to the lab concentrations (Figures 20 and 22; Figures 21 and 23).

In the laboratory concentrations for the DI water matrix, the Swiss labs kit consistently underestimated concentrations (Figure 2), while the citizen science data overestimated concentrations (Figures 20,22). Interestingly, in Amherst tap water, the Swiss Labs kit measured

0 mg/L in the laboratory-obtained data (Figure 2), but the citizen scientists did detect iron at a higher concentration of 12.3 mg/L (although it was underestimated as 5 mg/L) (Figures 21, 23).

While in the laboratory study, the Aquacheck kit for the DI water matrix had consistently performed well (Figure 3), when used by citizen scientists, the measurements were underestimates of the lab-obtained concentrations (Figures 20, 22). The Aquacheck laboratory-obtained data and citizen science data were very similar for the Amherst tap water matrix (Figures 3, 21, 23). Similarly, the Sensafe had overestimated concentrations in the DI water matrix in the lab portion of the study (Figure 3), but citizen scientists measured concentrations that were underestimates of the lab-obtained values (Figures 20, 22). In the laboratory-obtained data for the Amherst tap water matrix, the Sensafe kit had no detection while the citizen science data did have some detections at low concentrations (Figures 21, 23).

	The instructions were clear	The results were easy to read	I have confidence in the results	I know what these results mean for the safety of the water
<b>Sensafe</b> (EPA standard identified on bottle)	4.65	3.20	3.25	3.90
<b>Aquacheck</b> (add reducing agent)	4.58	4.46	4.21	3.17
<b>Swiss Labs</b> (combo kit)	5.00	3.67	3.22	3.44

**Figure 24.** Average results from surveys. Each column represents a statement and each row a test kit brand. Sensafe n = 20. Aquacheck n = 24. Swiss Labs n = 9. The highest average response for each question is highlighted in green. Citizen Scientists were asked to rate each statement on a scale of 1-5. 1 = strongly disagree, 3 = neutral, and 5 = strongly agree.

Citizen scientists were asked to rate statements about the test kits in order to provide insight on their satisfaction with each kit. The Swiss Labs kit were rated the highest for clear instructions (Figure 24). This was the only combination kit provided to citizen scientists. The Aquacheck kit was rated as the as easiest to read results and most confidence in the results (Figure 25). Notably, the Aquacheck kit was the only kit that required the addition of any

reagent, which may have contributed to users' feelings of confidence. The Sensafe kit was rated as highest for users to know what the results meant for the safety of water (Figure 25). Of the three kits given to citizen scientists, this was the only kit that identified what the EPA standard was on the bottle, which may have contributed to users feeling they understood the results. However, comparing the overall numbers, the highest rating for understanding safety of drinking water was 3.9 out of 5, which is low. Overall, the citizen scientists found the kit instructions clear but their ability to give meaningful and interpretable results lacking (Figure 25).

### 3.5. Discussion

The citizen science data provided more insight to the iron test kits, although interpretation of results were limited by small sample size. There were some variations between the citizen science results and the laboratory-obtained results. This could be due to sample size: there were fewer citizen science tests conducted per concentration compared to the laboratory-obtained data. In future experiments, it would be beneficial to increase the sample sizes per concentration in order to match the laboratory-obtained data, as well as include more test kits and more parameters in future experimentation. This could also be due to citizen scientists performing the methods or interpreting the results incorrectly.

In future work, it would be insightful to draft directions to account for a water sample that is over range of a test kit and what the next steps should be to get it in range and able to be measured by the kit. Then, from the data researchers could determine if citizen scientists could accurately deploy those methods.

Additionally, this research only included engineering undergraduate students; in the future it would be interesting to expand this audience such as younger children, college students studying non-science and technology related fields, or adults with no science backgrounds.

## Chapter 4: Conclusion

- Most of the evaluated test kits could not be used by citizen scientists to produce reliable scientific monitoring data in real water sources. As illustrated by the results and sensitivity analysis, the test strips most often do not produce accurate data, although some brands may be adequate for classification purposes (e.g. presence/absence of a contaminant, or above or below the SMCL). If citizen scientists want to collect data to monitor their drinking water at home or in the workplace, the test strips, as they are right now, are of varying accuracy, and it will depend on what else is in the water being tested. The majority of the test kits underestimated concentrations, potentially leading people to obtain false negative results.
- Procedures to eliminate reader error when matching colors on the test kits need to be implemented. Matching colors to a color chart has the potential to vary drastically from person to person and include factors such as time of day, lighting conditions, colorblindness, and eyesight. Implementing a procedure or additional method (e.g. a photograph) to consistently read these color charts could eliminate this large source of error.
- Citizen scientists can successfully use the test kits and could produce reliable data with reliable tools. Providing the test kits to citizen scientists within the UMass College of Engineering illustrated that they could successfully use the kits; although the accuracy of the kits and their instructions to users need further investigation. Therefore, if the kits themselves were improved, then they could be used to produce reliable monitoring data.

## References

- Assumpção, T. H., Popescu, I., Jonoski, A., & Solomatine, D. P. (2018). Citizen observations contributing to flood modelling: opportunities and challenges. *Hydrology and Earth System Sciences*, 22(2), 1473–1489. <https://doi.org/10.5194/hess-22-1473-2018>
- Au, J., Bagchi, P., Chen, B., Martinez, R., Dudley, S. A., & Sorger, G. J. (2000). Methodology for public monitoring of total coliforms, *Escherichia coli* and toxicity in waterways by Canadian high school students. *Journal of Environmental Management*, 58(3), 213–230. <https://doi.org/10.1006/jema.2000.0323>
- Barbara A. Israel, Amy J. Schulz, Edith A. Parker, & Becker, and A. B. (1998). REVIEW OF COMMUNITY-BASED RESEARCH: Assessing Partnership Approaches to Improve Public Health. *Annual Review of Public Health*, 19(1), 173–202. <https://doi.org/10.1146/annurev.publhealth.19.1.173>
- Baum, R., Bartram, J., & Hrudey, S. (2016). The Flint Water Crisis Confirms That U.S. Drinking Water Needs Improved Risk Management. *Environmental Science & Technology*, 50(11), 5436–5437. <https://doi.org/10.1021/acs.est.6b02238>
- Bird, T. J., Bates, A. E., Lefcheck, J. S., Hill, N. A., Thomson, R. J., Edgar, G. J., ... Frusher, S. (2014). Statistical solutions for error and bias in global citizen science datasets. *Biological Conservation*, 173, 144–154. <https://doi.org/10.1016/j.biocon.2013.07.037>
- Bonney, R., Cooper, C. B., Dickinson, J., Kelling, S., Phillips, T., Rosenberg, K. V., & Shirk, J. (2009). Citizen Science: A Developing Tool for Expanding Science Knowledge and Scientific Literacy. *BioScience*, 59(11), 977–984. <https://doi.org/10.1525/bio.2009.59.11.9>
- Bonney, R., Shirk, J. L., Phillips, T. B., Wiggins, A., Ballard, H. L., Miller-Rushing, A. J., & Parrish, J. K. (2014). Next steps for citizen science. *Science*, 343(6178), 1436–1437.

- Burgess, H. K., DeBey, L. B., Froehlich, H. E., Schmidt, N., Theobald, E. J., Ettinger, A. K., ... Parrish, J. K. (2017). The science of citizen science: Exploring barriers to use as a primary research tool. *Biological Conservation*, 208, 113–120.  
<https://doi.org/10.1016/j.biocon.2016.05.014>
- CDC. (2018, October). Copper and Drinking Water from Private Wells | Wells | Private Water Systems | Drinking Water | Healthy Water | CDC. Retrieved February 4, 2019, from <https://www.cdc.gov/healthywater/drinking/private/wells/disease/copper.html>
- Conrad, C. C., & Hilchey, K. G. (2011). A review of citizen science and community-based environmental monitoring: issues and opportunities. *Environmental Monitoring and Assessment*, 176(1–4), 273–291. <https://doi.org/10.1007/s10661-010-1582-5>
- Conrad, C. T., & Daoust, T. (2008). Community-Based Monitoring Frameworks: Increasing the Effectiveness of Environmental Stewardship. *Environmental Management*, 41(3), 358–366. <https://doi.org/10.1007/s00267-007-9042-x>
- Dickinson, J. L., Shirk, J., Bonter, D., Bonney, R., Crain, R. L., Martin, J., ... Purcell, K. (2012). The current state of citizen science as a tool for ecological research and public engagement. *Frontiers in Ecology and the Environment*, 10(6), 291–297.  
<https://doi.org/10.1890/110236>
- EPA. (2004). *Understanding the Safe Drinking Water Act* (No. EPA 816-F-04-030; p. 4). Office of Water (4606), Environmental Protection Agency.
- EPA. (2008). *Lead and Copper Rule: A Quick Reference Guide* (No. EPA 816-F-08-018; p. 2). Office of Water (4606M), Environmental Protection Agency.
- Farnham, D. J., Gibson, R. A., Hsueh, D. Y., McGillis, W. R., Culligan, P. J., Zain, N., & Buchanan, R. (2017). Citizen science-based water quality monitoring: Constructing a

- large database to characterize the impacts of combined sewer overflow in New York City. *Science of The Total Environment*, 580, 168–177.  
<https://doi.org/10.1016/j.scitotenv.2016.11.116>
- Farooqi, A., Masuda, H., & Firdous, N. (2007). Toxic fluoride and arsenic contaminated groundwater in the Lahore and Kasur districts, Punjab, Pakistan and possible contaminant sources. *Environmental Pollution*, 145(3), 839–849.  
<https://doi.org/10.1016/j.envpol.2006.05.007>
- Greenwood, R., Mills, G. A., & Roig, B. (2007). Introduction to emerging tools and their use in water monitoring. *TrAC Trends in Analytical Chemistry*, 26(4), 263–267.  
<https://doi.org/10.1016/j.trac.2006.12.006>
- Hochachka, W. M., Fink, D., Hutchinson, R. A., Sheldon, D., Wong, W.-K., & Kelling, S. (2012). Data-intensive science applied to broad-scale citizen science. *Trends in Ecology & Evolution*, 27(2), 130–137. <https://doi.org/10.1016/j.tree.2011.11.006>
- Khatri, N., Tyagi, S., & Rawtani, D. (2017). Recent strategies for the removal of iron from water: A review. *Journal of Water Process Engineering*, 19, 291–304.  
<https://doi.org/10.1016/j.jwpe.2017.08.015>
- MassDEP. (2018). *Private Well Guidelines* (p. 136).
- MassDEP. (2019a). Private Wells. Retrieved April 25, 2019, from Mass.gov website:  
<https://www.mass.gov/private-wells>
- MassDEP. (2019b). Protect Your Family: A Guide to Water Quality Testing for Private Wells. Retrieved April 25, 2019, from Mass.gov website: <https://www.mass.gov/service-details/protect-your-family-a-guide-to-water-quality-testing-for-private-wells>

- Metzger, M. (2005, January). *The Relationship Between Iron and pH*. National Testing Laboratories, Ltd.
- Miller-Rushing, A., Primack, R., & Bonney, R. (2012). The history of public participation in ecological research. *Frontiers in Ecology and the Environment*, *10*(6), 285–290. <https://doi.org/10.1890/110278>
- Newman, G., Graham, J., Crall, A., & Laituri, M. (2011). The art and science of multi-scale citizen science support. *Ecological Informatics*, *6*(3–4), 217–227. <https://doi.org/10.1016/j.ecoinf.2011.03.002>
- NSF. (2019a). Accreditations and Quality - NSF International. Retrieved April 25, 2019, from National Sanitation Foundation website: <http://www.nsf.org/about-nsf/accreditations>
- NSF. (2019b). Beverage and Packaged Ice Certification - NSF International. Retrieved April 25, 2019, from National Sanitation Foundation website: <http://www.nsf.org/services/by-industry/food-safety-quality/beverage-bottled-water/beverages-and-packaged-ice-certification>
- Ottinger, G. (2010). Buckets of Resistance: Standards and the Effectiveness of Citizen Science. *Science, Technology, & Human Values*, *35*(2), 244–270. <https://doi.org/10.1177/0162243909337121>
- Ottinger, G. (2017). Making sense of citizen science: Stories as a hermeneutic resource. *Energy Research & Social Science*, *31*, 41–49. <https://doi.org/10.1016/j.erss.2017.06.014>
- Pieper, K. J., Tang, M., & Edwards, M. A. (2017). Flint Water Crisis Caused By Interrupted Corrosion Control: Investigating “Ground Zero” Home. *Environmental Science & Technology*, *51*(4), 2007–2014. <https://doi.org/10.1021/acs.est.6b04034>

- Rawot, L. A., Peirce, J. J., & Kanoy, B. (2011). Preparing for the WATERS Network: Engaging Middle-School Science Students in Environmental Engineering. *Journal of Professional Issues in Engineering Education and Practice*, 137(3), 135–140.  
[https://doi.org/10.1061/\(ASCE\)EI.1943-5541.00000036](https://doi.org/10.1061/(ASCE)EI.1943-5541.00000036)
- Silvertown, J. (2009). A new dawn for citizen science. *Trends in Ecology & Evolution*, 24(9), 467–471.
- Sobsey, M. D., Stauber, C. E., Casanova, L. M., Brown, J. M., & Elliott, M. A. (2008). Point of Use Household Drinking Water Filtration: A Practical, Effective Solution for Providing Sustained Access to Safe Drinking Water in the Developing World. *Environmental Science & Technology*, 42(12), 4261–4267. <https://doi.org/10.1021/es702746n>
- Sullivan, B. L., Aycrigg, J. L., Barry, J. H., Bonney, R. E., Bruns, N., Cooper, C. B., ... Kelling, S. (2014). The eBird enterprise: An integrated approach to development and application of citizen science. *Biological Conservation*, 169, 31–40.  
<https://doi.org/10.1016/j.biocon.2013.11.003>
- Tulloch, A. I. T., Possingham, H. P., Joseph, L. N., Szabo, J., & Martin, T. G. (2013). Realising the full potential of citizen science monitoring programs. *Biological Conservation*, 165, 128–138. <https://doi.org/10.1016/j.biocon.2013.05.025>
- US EPA, O. (2015a, September 2). Secondary Drinking Water Standards: Guidance for Nuisance Chemicals [Overviews and Factsheets]. Retrieved January 30, 2019, from US EPA website: <https://www.epa.gov/dwstandardsregulations/secondary-drinking-water-standards-guidance-nuisance-chemicals>
- US EPA, O. (2015b, November 30). National Primary Drinking Water Regulations [Overviews and Factsheets]. Retrieved March 18, 2019, from US EPA website:

<https://www.epa.gov/ground-water-and-drinking-water/national-primary-drinking-water-regulations>

WHO. (2004). *fluoride in drinking water* (No. WHO/SDE/WSH/03.04/96). Retrieved from World Health Organization website:

[https://www.who.int/water\\_sanitation\\_health/dwq/chemicals/fluoride.pdf](https://www.who.int/water_sanitation_health/dwq/chemicals/fluoride.pdf)

WHO. (2011). *Manganese in Drinking-water* (No. WHO/SDE/WSH/03.04/104/Rev/1). Retrieved from World Health Organization website:

[https://www.who.int/water\\_sanitation\\_health/dwq/chemicals/manganese.pdf](https://www.who.int/water_sanitation_health/dwq/chemicals/manganese.pdf)

**APPENDIX**

**Table 14. Iron concentrations measured by test kits in DI Water. All tests were repeated five times.**

<b><u>Test Kit Brand</u></b>	<b><u>Measured Concentration (mg/L)</u></b>	<b><u>Mean OTS (mg/L)</u></b>	<b><u># &gt; 0 mg/L</u></b>	<b><u>Standard Deviation</u></b>
Aquacheck	0.0	0	0	0
Baldwin		0	0	0
Bartovention		0	0	0
JNW		0	0	0
Sensafe		0	0	0
Swiss Labs		0	0	0
Tech Pure		0	0	0
Water Works		0	0	0
Aquacheck	0.12	0	0	0
Baldwin		0	0	0
JNW		0	0	0
Sensafe		0.1	5	0
Water Works		0.3	5	0
Aquacheck	0.30	0.24	5	0.082
Water Works		0.38	5	0.110
Baldwin	0.33	0	0	0
JNW		0	0	0
Sensafe		1.35	5	0.602
Aquacheck	0.44	0.27	5	0.076
Water Works		0.90	5	0.224
Baldwin	0.55	0	0	0
JNW		0	0	0
Sensafe		1.55	5	0.622
Aquacheck	1.01	0.92	5	0.179
Bartovention		1	5	0
Swiss Labs		0	0	0
Tech Pure		0	0	0
Water Works		3	5	0
Baldwin		1.05	0	0
JNW	0		0	0
Sensafe	5		5	0
Aquacheck	1.91	1.6	5	0.548
Bartovention		2	5	0
Swiss Labs		1	5	0
Tech Pure		0.2	1	0.447
Water Works		3.8	5	1.10
Baldwin		3.10	0	0
JNW	5		5	0
Sensafe	5		5	0
Baldwin	4.42	0	0	0
JNW		5	5	0
Sensafe		5	5	0
Aquacheck	5.58	5	5	0
Bartovention		6	5	1.369
Swiss Labs		1.9	5	0.821
Tech Pure		3.2	5	1.643
Water Works		5	5	0
Baldwin		5.6	0	0

JNW		5	5	0
Sensafe		5	5	0
Bartovention	11	10	5	0
JNW		10	5	0
Swiss Labs		5	5	0
Tech Pure		5	5	0
Bartovention	23	25	5	0
JNW		25	5	0
Swiss Labs		10	5	0
Tech Pure		10	5	0
Bartovention	44	25	5	0
JNW		50	5	0
Swiss Labs		25	5	0
Tech Pure		25	5	0
Bartovention	107	25	5	0
JNW		100	5	0
Swiss Labs		100	5	0
Tech Pure		100	5	0

Table 15. Iron concentrations measured by test kits in Amherst Tap Water. All tests were repeated five times.

<b><u>Test Kit Brand</u></b>	<b><u>Measured Concentration (mg/L)</u></b>	<b><u>Mean OTS (mg/L)</u></b>	<b><u># &gt; 0 mg/L</u></b>	<b><u>Standard Deviation</u></b>
Sensafe	0.14	0.02	5	0
Aquacheck	0.24	0.15	5	0
Baldwin		0	0	0
Bartovention		0	0	0
JNW		0	0	0
Swiss Labs		0	0	0
Tech Pure		0	0	0
Water Works		0.3	5	0
Sensafe	0.33	0.02	5	0
Aquacheck	0.39	0.3	5	0
Baldwin		0	0	0
Bartovention		0	0	0
JNW		0	0	0
Swiss Labs		0	0	0
Tech Pure		0	0	0
Water Works		1	5	0
Sensafe	0.53	0.038	5	0.0164
Aquacheck	0.60	0.60	5	0
Baldwin		0	0	0
Bartovention		0	0	0
JNW		0	0	0
Swiss Labs		0	0	0
Tech Pure		0	0	0
Water Works		1	5	0
Sensafe	0.93	0.05	5	0
Aquacheck	1.17	0.60	5	0
Baldwin		0	0	0
Bartovention		0	0	0
JNW		0	0	0

Swiss Labs		0	0	0
Tech Pure		0	0	0
Water Works		3	5	0
Sensafe	1.89	0.032	5	0.0164
Aquacheck	3.30	3.8	5	1.643
Baldwin		0	0	0
Bartovention		0	0	0
JNW		0	0	0
Swiss Labs		0	0	0
Tech Pure		0	0	0
Water Works		5	5	0
Aquacheck	4.22	5	5	0
Baldwin		0	0	0
Bartovention		0	0	0
JNW		0	0	0
Swiss Labs		0	0	0
Tech Pure		0	0	0
Water Works		5	5	0
Sensafe	4.54	0.032	5	0.0164
Aquacheck	4.84	5	5	0
Baldwin		0	0	0
Bartovention		0	0	0
JNW		0	0	0
Swiss Labs		0	0	0
Tech Pure		0	0	0
Water Works		5	5	0
Baldwin	7.60	0	0	0
Bartovention		2	5	0
JNW		0	0	0
Swiss Labs		0	0	0
Tech Pure		0	0	0
Baldwin	22.5	10	5	0
Bartovention		10	5	0
JNW		10	5	0
Swiss Labs		5	5	0
Tech Pure		8	5	2.739
Baldwin	48	50	5	0
Bartovention		50	5	0
JNW		25	5	0
Swiss Labs		25	5	0
Tech Pure		25	5	0
Baldwin	97	100	5	0
Bartovention		50	5	0
JNW		100	5	0
Swiss Labs		100	5	0
Tech Pure		50	5	0

Table 16. Iron concentrations measured by test kits in Mill River Water. All tests were repeated five times.

<u>Test Kit Brand</u>	<u>Measured Concentration (mg/L)</u>	<u>Mean OTS (mg/L)</u>	<u># &gt; 0 mg/L</u>	<u>Standard Deviation</u>
Sensafe	0.06	0.016	4	0.009

Aquacheck	0.20	0.15	5	0
Baldwin		0	0	0
Bartovention		0	0	0
JNW		0	0	0
Swiss Labs		0	0	0
Tech Pure		0	0	0
Water Works		0.3	5	0
Sensafe	0.33	0.026	5	0.013
Aquacheck	0.36	0.24	5	0.082
Baldwin		0	0	0
Bartovention		0	0	0
JNW		0	0	0
Swiss Labs		0	0	0
Tech Pure		0	0	0
Water Works		0.5	5	0
Aquacheck	0.53	1	5	0
Baldwin		0	0	0
Bartovention		0	0	0
JNW		0	0	0
Swiss Labs		0	0	0
Tech Pure		0	0	0
Water Works		0.54	5	0.134
Sensafe	0.67	0.050	5	0
Sensafe	0.85	0.044	5	0.013
Aquacheck	0.94	0.68	5	0.179
Baldwin		0	0	0
Bartovention		0	0	0
JNW		0	0	0
Swiss Labs		0	0	0
Tech Pure		0	0	0
Water Works		3	5	0
Sensafe	1.79	0.18	5	0.045
Aquacheck	3.04	5	5	0
Baldwin		0	0	0
Bartovention		0	0	0
JNW		0	0	0
Swiss Labs		0	0	0
Tech Pure		0	0	0
Water Works		5	5	0
Aquacheck	3.86	5	5	0
Baldwin		0	0	0
Bartovention		1	5	0
JNW		0	0	0
Swiss Labs		0	0	0
Tech Pure		0	0	0
Water Works		5	5	0
Sensafe	4.24	0.90	5	0.137
Aquacheck	4.90	5	5	0
Baldwin		0	0	0
Bartovention		2.6	5	0.548
JNW		0	0	0
Swiss Labs		0	0	0
Tech Pure		0	0	0
Water Works		5	5	0
Baldwin	7.5	5	5	0
Bartovention		5	5	0

JNW		5	5	0
Swiss Labs		0	0	0
Tech Pure		0	0	0
Baldwin	23	25	5	0
Bartovention		25	5	0
JNW		25	5	0
Swiss Labs		10	5	0
Tech Pure		10	5	0
Baldwin	42	50	5	0
Bartovention		50	5	0
JNW		50	5	0
Swiss Labs		50	5	0
Tech Pure		25	5	0
Baldwin	104	100	5	0
Bartovention		50	5	0
JNW		100	5	0
Swiss Labs		100	5	0
Tech Pure		50	5	0

Table 17. Copper concentrations measured by test kits in DI Water. All tests were repeated five times.

<b>Test Kit Brand</b>	<b>Measured Concentration (mg/L)</b>	<b>Mean OTS (mg/L)</b>	<b># &gt; 0 mg/L</b>	<b>Standard Deviation</b>
Sensafe	0.0	0	0	0
Sensafe	0.09	0.06	5	0.0224
Tech Pure	0.19	0.4	2	0.548
Swiss Labs		0	0	0
JNW		0.1	2	0.137
Baldwin		0.2	2	0.274
Aquacheck		0.22	4	0.179
Sensafe	0.27	0.09	5	0.0224
Sensafe	0.37	0.20	5	0
Tech Pure	0.39	0.6	3	0.548
Swiss Labs		0	0	0
JNW		0.05	1	0.112
Baldwin		0.2	2	0.274
Aquacheck		0.28	4	0.217
Sensafe	0.88	0.88	5	0.268
Tech Pure	0.98	0	0	0
Swiss Labs		0.5	5	0
Baldwin		0.9	5	0.224
Aquacheck		0	0	0
Sensafe	1.83	2	5	0
Tech Pure	2.75	0	0	0
Swiss Labs		1	5	0
Baldwin		1.7	5	1.204
Aquacheck		3	5	0
Tech Pure	9.20	4.6	5	4.93
Swiss Labs		10	5	0
Baldwin		6.7	5	4.60

Table 18. Copper concentrations measured by test kits in Amherst, MA Tap Water. All tests were repeated five times.

<u>Test Kit Brand</u>	<u>Measured Concentration (mg/L)</u>	<u>Mean OTS (mg/L)</u>	<u># &gt; 0 mg/L</u>	<u>Standard Deviation</u>
Baldwin	0.10	0.05	5	0
JNW		0.50	5	0
Sensafe	0.18	0.24	5	0.09
Aquacheck	0.23	0.08	2	0.110
Swiss Labs		0	0	0
Tech Pure		0	0	0
Baldwin	0.30	0.05	5	0
JNW		0.50	5	0
Baldwin	0.48	0.05	5	0
JNW		0.50	5	0
Sensafe	0.54	0.60	5	0.374
Aquacheck	0.63	0.38	5	0.164
Swiss Labs		0	0	0
Tech Pure		0	0	0
Baldwin		0.05	5	0
JNW	1.01	0.50	5	0
Sensafe		1.60	5	0.548
Aquacheck	1.15	3	5	0
Swiss Labs		0.4	4	0.224
Tech Pure		0	0	0
Baldwin		0.05	5	0
JNW	2.08	1.0	5	0
Sensafe		2	5	0
Aquacheck	3.09	3	5	0
Swiss Labs		2.20	5	0.110
Tech Pure		0	0	0
Aquacheck		0	0	0
Swiss Labs	8.72	10	5	0
Tech Pure		1	5	0

Table 19. Copper concentrations measured by test kits in Mill River Water. All tests were repeated five times.

<u>Test Kit Brand</u>	<u>Measured Concentration (mg/L)</u>	<u>Mean OTS (mg/L)</u>	<u># &gt; 0 mg/L</u>	<u>Standard Deviation</u>
Sensafe	0.09	0.07	5	0.027
Baldwin	0.10	0.05	5	0
JNW		0	0	0
Aquacheck	0.12	0	0	0
Swiss Labs		0	0	0
Tech Pure		0	0	0
Baldwin		0.05	5	0

JNW		0	0	0
Sensafe	0.15	0.08	5	0.027
Sensafe	0.31	0.24	5	0.089
Baldwin	0.37	0.05	5	0.089
JNW		0	0	0
Aquacheck	0.42	0.32	5	0.164
Swiss Labs		0	0	0
Tech Pure		0	0	0
Sensafe	0.77	0.36	5	0.089
Baldwin	0.83	0.05	5	0
JNW		0.50	5	0
Aquacheck	0.90	0.5	5	0
Swiss Labs		0	0	0
Tech Pure		0	0	0
Sensafe	1.51	1.8	5	0.447
Baldwin	1.92	0.05	5	0
JNW		1	5	0
Aquacheck	2.70	2.6	5	0.894
Swiss Labs		1	5	0
Tech Pure		0	0	0
Aquacheck	8.72	0	0	0
Swiss Labs		10	5	0
Tech Pure		1	5	0

Table 20. Fluoride concentrations measured by test kits in DI Water. All tests were repeated five times.

<b>Test Kit Brand</b>	<b>Measured Concentration (mg/L)</b>	<b>Mean OTS (mg/L)</b>	<b># &gt; 0 mg/L</b>	<b>Standard Deviation</b>
Baldwin	0	0	0	0
JNW		0	0	0
Swiss Labs		0	0	0
Tech Pure		0	0	0
Baldwin	0.52	0	0	0
JNW		0	0	0
Swiss Labs		0	0	0
Tech Pure		0.1	1	0.224
Baldwin	0.87	0	0	0
JNW		0	0	0
Swiss Labs		0	0	0
Tech Pure		0.5	5	0
Baldwin	1.70	0	0	0
JNW		0	0	0
Swiss Labs		0	0	0
Tech Pure		0.412	5	0.204
Baldwin	3.40	0	0	0
JNW		0	0	0
Swiss Labs		0	0	0
Tech Pure		3.2	5	1.643
Baldwin	3.80	0	0	0
JNW		0	0	0
Swiss Labs		0	0	0
Tech Pure		3.2	5	1.643
Baldwin	10.60	0	0	0

JNW		0	0	0
Swiss Labs		0	0	0
Tech Pure		3.8	5	1.643
Baldwin	36	0.5	5	0
JNW		0	0	0
Swiss Labs		6.0	2	5.477
Baldwin	53	0.5	5	0
JNW		0	0	0
Swiss Labs		6	3	5.477
Baldwin	179	4.6	5	0.5477
JNW		10	5	0
Swiss Labs		100	5	0

Table 21. Fluoride concentrations measured by test kits in Amherst Tap Water. All tests were repeated five times.

<u>Test Kit Brand</u>	<u>Measured Concentration (mg/L)</u>	<u>Mean OTS (mg/L)</u>	<u># &gt; 0 mg/L</u>	<u>Standard Deviation</u>
Baldwin	1.14	0.10	1	0.224
JNW		0	0	0
Swiss Labs		0	0	0
Tech Pure		0.40	4	0.224
Baldwin	1.85	0	0	0
JNW		0	0	0
Swiss Labs		0	0	0
Tech Pure		0.7	5	0.274
Baldwin	2.0	0	0	0
JNW		0	0	0
Swiss Labs		0	0	0
Tech Pure		0	2	0.274
Baldwin	4.60	0.10	2	0.211
JNW		0	0	0
Swiss Labs		0	0	0
Tech Pure		0.45	5	0.497
Baldwin	15.8	0.3	3	0.279
JNW		0	0	0
Swiss Labs		0	0	0
Tech Pure		1.2	5	0.447
Baldwin	26	0.8	5	0.279
JNW		0	0	0
Swiss Labs		4.0	2	5.477
Tech Pure		1.8	5	0.447
Baldwin	58	1	5	0
JNW		0	0	0
Swiss Labs		25	5	0
Tech Pure		3.80	5	1.643

Table 22. Fluoride concentrations measured by test kits in Mill River Water. All tests were repeated five times.

<u>Test Kit Brand</u>	<u>Measured Concentration (mg/L)</u>	<u>Mean OTS (mg/L)</u>	<u># &gt; 0 mg/L</u>	<u>Standard Deviation</u>
Baldwin	0.57	0	0	0
JNW		0	0	0
Swiss Labs		0	0	0
Tech Pure		0.6	5	0.223
Baldwin	1.46	0	0	0
JNW		0	0	0
Swiss Labs		0	0	0
Tech Pure		0.4	2	0.548
Baldwin	3.80	0.1	1	0.224
JNW		0	0	0
Swiss Labs		0	0	0
Tech Pure		0.6	5	0.224
Baldwin	4.60	0	0	0
JNW		0	0	0
Swiss Labs		0	0	0
Tech Pure		0.7	5	0.274
Baldwin	5.60	0	0	0
JNW		0	0	0
Swiss Labs		0	0	0
Tech Pure		0.7	5	0.274
Baldwin	17	2	1	4.47
JNW		0	0	0
Swiss Labs		2	1	4.47
Tech Pure		0.7	4	0.447
Baldwin	20	0.8	5	0.274
JNW		0	0	0
Swiss Labs		4.0	2	5.477
Tech Pure		1.8	5	0.447
Baldwin	59	1	5	0
JNW		0	0	0
Swiss Labs		25	5	0
Tech Pure		4.4	5	1.341

Table 23. Manganese concentrations measured by test kits in DI Water. All tests were repeated five times.

<u>Test Kit Brand</u>	<u>Measured Concentration (mg/L)</u>	<u>Mean OTS (mg/L)</u>	<u># &gt; 0 mg/L</u>	<u>Standard Deviation</u>
Sensafe	0	<0.02	5	0
Sensafe	0.044	0.02	5	0
Sensafe	0.099	0.08	5	0.027
Sensafe	0.20	0.09	5	0.022
Sensafe	2.45	>1.6	5	0

Table 24. Manganese concentrations measured by test kits in Amherst Tap Water. All tests were repeated five times.

<u>Test Kit Brand</u>	<u>Measured Concentration (mg/L)</u>	<u>Mean OTS (mg/L)</u>	<u># &gt; 0 mg/L</u>	<u>Standard Deviation</u>
Sensafe	0.009	0.06	5	0.022
Sensafe	0.057	0.12	5	0.045
Sensafe	0.108	0.09	5	0.022
Sensafe	0.404	0.36	5	0.089
Sensafe	7.70	>1.6	5	0

Table 25. Manganese concentrations measured by test kits in Mill River Water. All tests were repeated five times.

<u>Test Kit Brand</u>	<u>Measured Concentration (mg/L)</u>	<u>Mean OTS (mg/L)</u>	<u># &gt; 0 mg/L</u>	<u>Standard Deviation</u>
Sensafe	0.008	0.08	5	0.027
Sensafe	0.062	0.12	5	0.045
Sensafe	0.111	0.12	5	0.045
Sensafe	0.394	0.32	5	0.110
Sensafe	8.2	>1.6	5	0