July 2019

"WAITING TO DIE"

Toxic Emissions and Disease Near the Louisiana Denka / DuPont Plant

UNIVERSITY NETWORK FOR HUMAN RIGHTS

Suggested citation:
University Network for Human Rights. *"Waiting to Die:" Toxic Emissions and Disease Near the Louisiana Denka/DuPont Plant.* July 2019. Retrieved from: http://www.humanrightsnetwork.org/waiting-to-die.

Please direct all inquiries to:
ruhan@humanrightsnetwork.org

CONTENTS

Executive Summary

Executive Summary

Three years ago, residents living near a chemical plant in St. John the Baptist Parish, Louisiana were told by the Environmental Protection Agency (EPA) that they faced the highest risk in the country of developing cancer from air pollution.

St. John Parish is part of an area of Louisiana known as "Cancer Alley," an 85-mile stretch of land along the Mississippi River between New Orleans and Baton Rouge. More than 150 chemical plants and oil refineries dot this stretch of land, where most communities are predominantly Black and many residents attribute seemingly staggering levels of cancer and other illness to toxic air emissions from industry.[1]

The St. John plant's neoprene manufacturing unit—owned by DuPont until its sale to Japanese company Denka Performance Elastomer[2] in November 2015—has been pumping the toxic chemical chloroprene into a predominantly Black community since 1969. Residents had long felt that there was too much illness in the area—far beyond what could be considered normal. One resident with whom we spoke recalled the words of her niece, shortly before she passed away of cancer: "We're just sitting here, waiting to die."

The EPA's 2011 National Air Toxics Assessment (NATA), released in December 2015, seemed to confirm many residents' suspicions. According to the most recent NATA, the risk of developing cancer from air pollution in the census tract closest to the Denka neoprene facility is nearly 50 times the national average due to emissions of chloroprene,[3] classified by the EPA as a "likely human carcinogen." The EPA advocates a significant reduction in chloroprene emissions from the Denka facility, such that air concentration of the chemical does not exceed 0.2 micrograms per cubic meter ($\mu g/m^3$)—the maximum concentration that would keep cancer risk from air pollution within the EPA's "upper limit of acceptability."[4]

For the past three years, community members have demanded a reduction of chloroprene emissions to this EPA-recommended maximum level of 0.2 $\mu g/m^3$. Their struggle for environmental justice has gained increasing traction and national me-

1 Trymaine Lee, *Cancer Alley: Big Industry, Big Problems*, MSNBC, http://www.msnbc.com/interactives/geography-of-poverty/se.html.

2 Denka Co. Ltd. owns 70% of Denka Performance Elastomer, and Mitsui Co. Ltd. owns 30%.

3 United States Environmental Protection Agency, 2014 National Air Toxics Assessment, https://www.epa.gov/national-air-toxics-assessment/2014-national-air-toxics-assessment.

4 Memorandum from Kelly Rimer to Frances Verhalen, Preliminary Risk-Based Concentration Value for Chloroprene in Ambient Air, United States Environmental Protection Agency (May 5, 2016), https://www.epa.gov/sites/production/files/2016-06/documents/memo-prelim-risk-based-concentrations050516.pdf.

dia coverage.[5]

Although the EPA has affirmed its confidence in the scientific validity of its chloroprene assessment—stating that the assessment "was developed using a robust, transparent, and public process and represents the Agency's top tier source of toxicity information on chloroprene"[6]—Denka continues to challenge the EPA's findings on chloroprene toxicity.[7]

Denka signed a voluntary agreement to reduce emissions in January 2017 and finished installing emissions reduction technology by the end of that year, but the EPA's air monitoring data continue to show high levels of chloroprene emissions—well in excess of the 0.2 μg/m³ guideline—in the neighborhoods around the Denka facility.

This report by the University Network for Human Rights presents localized health data from the area surrounding the Denka/DuPont plant. In March 2018, a team of trained researchers collected health data from a large sample of residents who live within 2.5 kilometers of the plant. Below, we present our in-depth analysis of this household health survey data. Our data reveal **extremely improbable rates of cancer and other illness among residents surveyed.** We also found that **prevalence of cancer and other illness among our survey sample is correlated with proximity to the Denka plant**, with higher rates of illness closer to the plant.

Cancer prevalence among those surveyed is unusually high. Among respondents (those who provided health information about themselves and all their household members), the p-value for cancer prevalence when compared to a distribution of populations with the same race, sex, and age demographics is **0.6% (very statistically significant)**.[8] In other words, the probability of the 9.7% cancer prevalence outcome that we found among respondents—the likelihood that we would see a cancer prevalence this high or higher in a population with the same race, sex, and age composition—is only 0.6%.

Among all residents surveyed (respondents plus all their household members, i.e. everyone for whom we collected information), the p-value for cancer prevalence is **3.43% (statistically significant)**.

---

5 See, e.g., Sharon Lerner, *When Pollution is a Matter of Life and Death*, New York Times (June 22, 2019), https://www.nytimes.com/2019/06/22/opinion/sunday/epa-carniogens.html; Jamiles Lartey and Oliver Laughland, *'Almost every household has someone that has died from cancer,'* The Guardian (May 6, 2019), https://www.theguardian.com/us-news/ng-interactive/2019/may/06/cancertown-louisana-reserve-special-report; Rebecca Hersher, *After Decades of Air Pollution, A Louisiana Town Rebels Against A Chemical Giant*, NPR (Mar. 6, 2018), https://www.npr.org/sections/health-shots/2018/03/06/583973428/after-decades-of-air-pollution-a-louisiana-town-rebels-against-a-chemical-giant; Victor Blackwell et al., *Toxic tensions in the heart of 'Cancer Alley,'* CNN (Oct. 20, 2017), https://www.cnn.com/2017/10/20/health/louisiana-toxic-town/index.html.

6 Memorandum from John Vandenberg to Wren Stenger, *EPA's Integrated Risk Information System (IRIS) Assessment of Chloroprene,* United States Environmental Protection Agency (May 25, 2016), https://www.epa.gov/sites/production/files/2016-06/documents/memo-iris-chloroprene052516.pdf.

7 Letter from Koki Tabuchi to Scott Pruitt, *Request to Withdraw and Correct the 2010 IRIS Review of Chloroprene* (June 26, 2017), https://www.scribd.com/embeds/408326975/content?start_page=1&view_mode=scroll&access_key=key-pMd8zNOfLwIOotdRyQXN&show_recommendations=false.

8 P-values are generally expressed as decimals rather than percentages (0.006 rather than 0.6%, for example). Throughout this report, we express p-values as percentages because they are conceptually easier for the layperson to understand this way. P-values less than 5% (0.05) are considered statistically significant.

Cancer prevalence among those surveyed is also associated strongly with proximity to the Denka facility. Cancer prevalence among respondents who live closest to the facility (within 1.5 kilometers) is 71% higher than the national rate, with a p-value of **0.26% (very statistically significant)**. Cancer prevalence among residents who live closest to the facility (within 1.5 kilometers) is 44% higher than the national rate, with a p-value of **0.33% (very statistically significant)**.

Prevalence of non-cancer health conditions associated with chloroprene exposure is also striking and invariably correlated with proximity to the plant. **Nearly half the children** in the households surveyed within 1.5 kilometers of the plant suffer from **headaches, nosebleeds, or both**. P-values for tachycardia (abnormally fast heart rate) diagnosed by a doctor or other health care provider are 0% for both respondents and residents, indicating **a virtual impossibility that high tachycardia prevalence among the survey sample was due to chance**.

Among respondents surveyed within 1.5 kilometers of the plant: **nearly 40%** regularly experience chest pain, heart palpitations, or both; **one-third** regularly experience wheezing and/or difficulty breathing; **more than half** regularly experience headaches, dizziness, and/or lightheadedness; **nearly half** regularly experience eye pain/irritation and/or watery eyes; **more than 40%** experience cough, sneezing, and/or sore/hoarse throat most of the time; **more than one-third** regularly experience skin rash/irritation and/or itchy skin; and **nearly 30%** experience fatigue/lethargy most of the time.

Overall, our findings strongly indicate that prevalence of cancer and other illness among residents surveyed is unusually high compared to what we would expect using national actuarial tables. **These results are disturbing enough to warrant additional in-depth, localized, and rigorous health studies in the area surrounding the Denka/DuPont plant and throughout Cancer Alley.**

In the meantime, local, state, and federal agencies—including the Louisiana Department of Environmental Quality and the Louisiana Department of Health—must insist that Denka Performance Elastomer adhere to the EPA's 0.2 $\mu g/m^3$ guideline for maximum chloroprene air concentration.

Study Methodology

Study Methodology

In March 2018, a team of trained researchers conducted a household health survey of the area surrounding the Denka neoprene plant in St. John the Baptist Parish, Louisiana. The University Network for Human Rights coordinated subsequent phases of the study, including data analysis and the production of this report. The purpose of the study was to determine the overall health status of a large sample of residents living in the area of the Denka facility and evaluate the relationship, if any, between chloroprene emissions from the facility and illness among area residents.

Epidemiology and statistics experts at Stanford University provided input and guidance to ensure use of proper actuarial processes, study design methods, and survey implementation principles and techniques. After undergoing intensive training and practice in survey implementation, fourteen Stanford University undergraduates implemented the survey over a period of 9 days. Our study methodology at each stage of the project is described in detail below.

SURVEY INSTRUMENT

Validated measures were used to formulate as many of the survey instrument questions as possible. The instrument was designed to collect certain health and other information—including age, sex, part- or full-time residency status, cancer and other medical diagnoses, and child health—about all residents of a household. Additional information was collected about respondents (those who took the survey) only, including race/ethnicity and medical symptoms.

Many symptoms and diagnoses were included in the survey instrument because of their link to chloroprene exposure, according to scientific literature. Other symptoms and diagnoses were included after residents of the area identified them as particular sources of concern in focus group sessions that our team held in February 2018.

After piloting a draft survey instrument with five residents of the area in February 2018, we modified the structure and wording of some questions for clarity and efficiency.

STUDY DESIGN

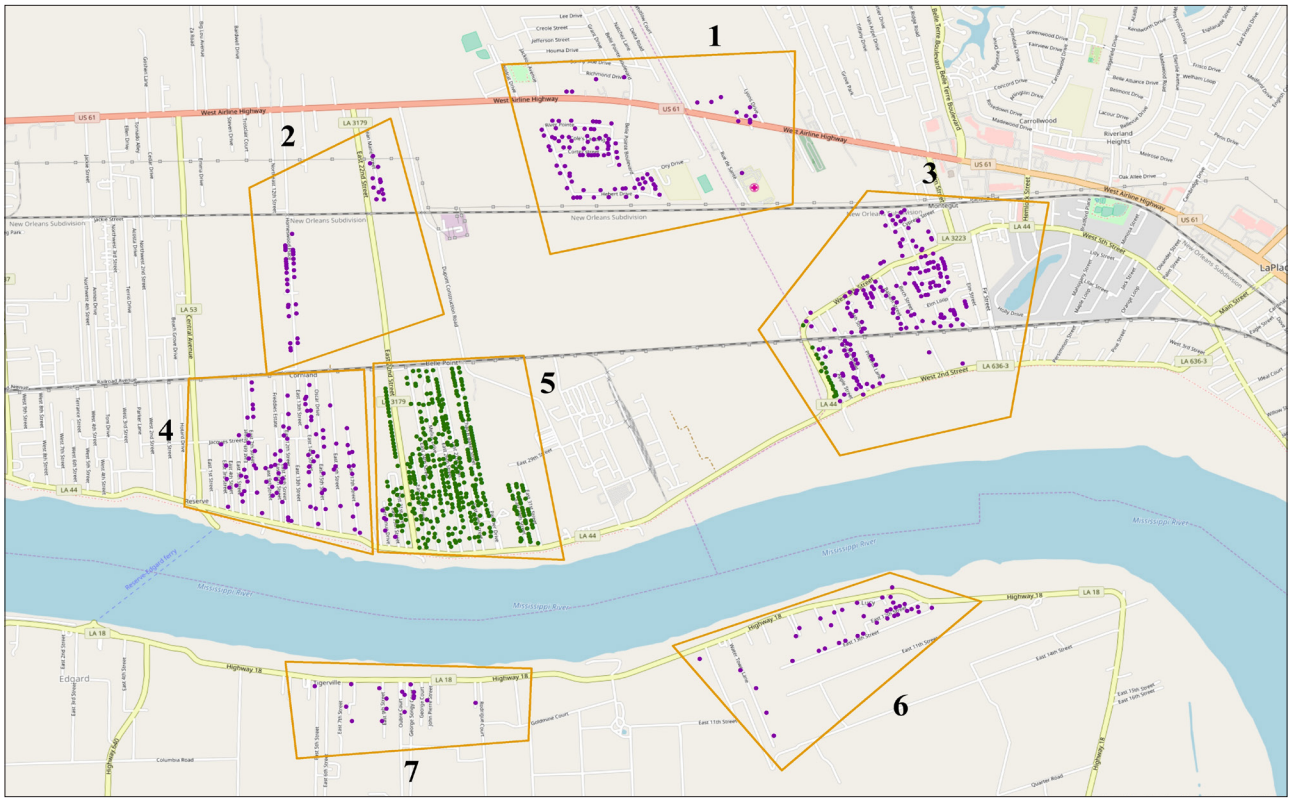
The geographic scope of the study was the area within a 2.5-kilometer radius of the Denka facility. In the images on the next page, the outer circle circumscribes the entire survey area and the inner circle circumscribes the area within 1.5 kilometers of the facility. The facility—with a red dot at its center—can be seen at the center of the survey area.

Maps of the survey area. The outer circle circumscribes the area within a 2.5-km radius of the Denka facility. The inner circle circumscribes the area within a 1.5-km radius of the Denka facility. The facility can be seen at the center of the survey area. In the map on the right, the grey dots represent households and each color represents a different census tract. Residents of the orange-colored census tract face the highest risk in the country of developing cancer from air pollution, according to the EPA.

Consistent with our hypothesis that rates of cancer and other illness increase with proximity to the plant, we designed our study protocol so that households within the 1.5-kilometer radius of the plant ("Zone 1," as shown in the images above) were surveyed at a higher proportion than households located between 1.5 and 2.5 kilometers from the plant ("Zone 2").

After obtaining addresses by census block online, we used a census batch geocoder to geocode the addresses. We determined that there are 445 total households in Zone 1 and 1,376 total households in Zone 2, according to 2010 census information. We designed our protocol to ensure that we would randomly survey at least 250 households in Zone 1 (56% of the Zone 1 total) and at least 250 households in Zone 2 (18% of the Zone 2 total). Assuming a survey response rate of approximately 50%, we used the R random number generator to generate a randomly-ordered list of all 445 addresses in Zone 1 (predicting that we would need to attempt to survey all 445 households to achieve our target number of 250 surveys in Zone 1). We also used the R random number generator to randomly select (and randomly order) 500 addresses in Zone 2 (predicting that we would need to attempt to survey at least 500 households to achieve our target number of 250 surveys in Zone 2).

We then divided the survey area into 7 geographic sub-areas for ease of survey implementation (that is, so that survey implementers could be assigned to a sub-area for a given period of time rather than having to walk long distances from household to household across the entire survey area):



In the above image, the green dots represent the 445 households on our Zone 1 list and the purple dots represent the 500 addresses on our Zone 2 list. We arbitrarily labeled each of the 7 geographic sub-areas (which are unrelated to the Zone 1/Zone 2 distinction) 1-7, as shown. (Note that two of the sub-areas, 3 and 5, contain a mix of Zone 1 and Zone 2 households. Sub-area 3 contains mostly households in Zone 2 and a few households in Zone 1, and sub-area 5 contains mostly households in Zone 1 and a few households in Zone 2.)

We then separated the addresses on the Zone 1 and Zone 2 lists by geographic sub-area, preserving the random order of the addresses within each sub-area. The number of addresses in each sub-area, and the percentage of the Zone 1 or Zone 2 list that each number represented, were as follows:

	Zone 1	**Zone 2**
sub-area 1		95 (19%)
sub-area 2		44 (8.8%)
sub-area 3	18 (4%)	201 (40.2%)
sub-area 4		99 (19.8%)
sub-area 5	427 (96%)	5 (1%)
sub-area 6		39 (7.8%)
sub-area 7		17 (3.4%)
Total	**445 (100%)**	**500 (100%)**

To calculate the number of households we needed to survey in each geographic sub-area, we multiplied each percentage in the table above by 250 (the target number of surveys for each of the two zones), rounding decimals up to the nearest whole number:

Target number of surveys for each sub-area

	Zone 1	Zone 2
sub-area 1		47.5 → 48
sub-area 2		22
sub-area 3	10	100.5 → 101
sub-area 4		49.5 → 50
sub-area 5	240	2.5 → 3
sub-area 6		19.5 → 20
sub-area 7		8.5 → 9
Total	**250**	**250 → 253**

Once we had attempted to survey all 500 addresses on our Zone 2 list at least twice without reaching the target number of surveys (250), we generated a randomly-ordered list of all remaining households in Zone 2 and separated the addresses by geographic sub-area. To reach the target number of surveys in each geographic sub-area (as listed in the table above), we attempted to survey almost every household in Zone 2 and every household in Zone 1. Thus, the survey response rate for each zone is equivalent to the percentage of households surveyed in each zone. **We ultimately surveyed a total of 267 (out of 445) households in Zone 1 (60%) and 271 households (out of 1,376) in Zone 2 (20%).**

STUDY PROTOCOL

One day prior to the start of survey implementation, we distributed flyers throughout the survey area. The flyers informed residents about the upcoming health survey, its purpose, and the possibility that their household might be randomly selected for participation. The flyers also stated that residents' participation in the survey was entirely voluntary.

After undergoing intensive training and practice in survey implementation principles and techniques under the supervision of Stanford University experts, a team of 14 Stanford undergraduates implemented the survey over 9 days (March 22-30, 2018). Households were surveyed from approximately 9am to 7pm each day.

Survey implementers almost always worked in pairs. Each day, we assigned each pair of survey implementers to one of the 7 geographic sub-areas and provided them with a list of household addresses in their sub-area. The list was randomized (see *Study Design*, above), but we optimized the route efficiency for each set of 20 addresses (using https://www.routexl.com) to reduce time spent walking between households. Survey implementers attempted to survey each of the 20 route-optimized households twice before moving on to the next set of 20. The following day, survey implementers made

a third attempt to survey households that had been attempted twice the previous day, before moving on to the next set of households. Survey implementers generally did not visit a household more than three times. If a household member declined to participate in the survey, implementers did not attempt to survey that household again.

Upon encountering a potential respondent, survey implementers introduced themselves and conveyed the purpose of the survey. They explained that participation in the survey was voluntary; that, if the potential respondent chose to participate, neither their name nor the names of any of their household members would be recorded; that any information provided would remain strictly confidential and would not be shared outside our research team; and that the overall results of the study would be made public but no one's identity or identifying health information would be disclosed. If the respondent verbally consented to participate in the survey, one of the survey implementers asked the survey questions while the other recorded the respondent's answers on a paper survey.

For each household surveyed, one household member (the "respondent") provided health and demographic information about themself and every other person living in the household. We use the term "residents" to refer to everyone for whom data were collected (that is, respondents plus all other household members).

Following completion of survey implementation, the data from each survey were manually entered into an electronic REDCap instrument.

LIMITATIONS

The most significant limitation of this study is our reliance on a single household member, the respondent, to provide health information about all other members of the household. It appeared that respondents did not always know the full extent of their fellow household members' medical diagnoses (with the exception of their non-adult children) and, as a result, may have underreported the diagnoses of others.[1] We believe that this limitation accounts for the discrepancy between rates of illness documented among "respondents" (those who took the survey) versus "residents" (respondents plus their household members). Overall, we documented higher levels of illness among respondents than among residents, and we attribute this to the fact that respondents generally appeared to be more knowledgeable about their own health than about the health of their household members (with the exception of their non-adult children). Thus, we believe that our resident health data are undercounts and that actual levels of illness among residents surveyed may be higher than the levels we documented.

In addition, community members often spoke to us about stigma associated with illness—especially cancer—among their family members, friends, and neighbors. We were told that many people "hide" their illnesses and suffering from others. This stigma may have manifested as a nonresponse bias that favored healthier individuals and households.

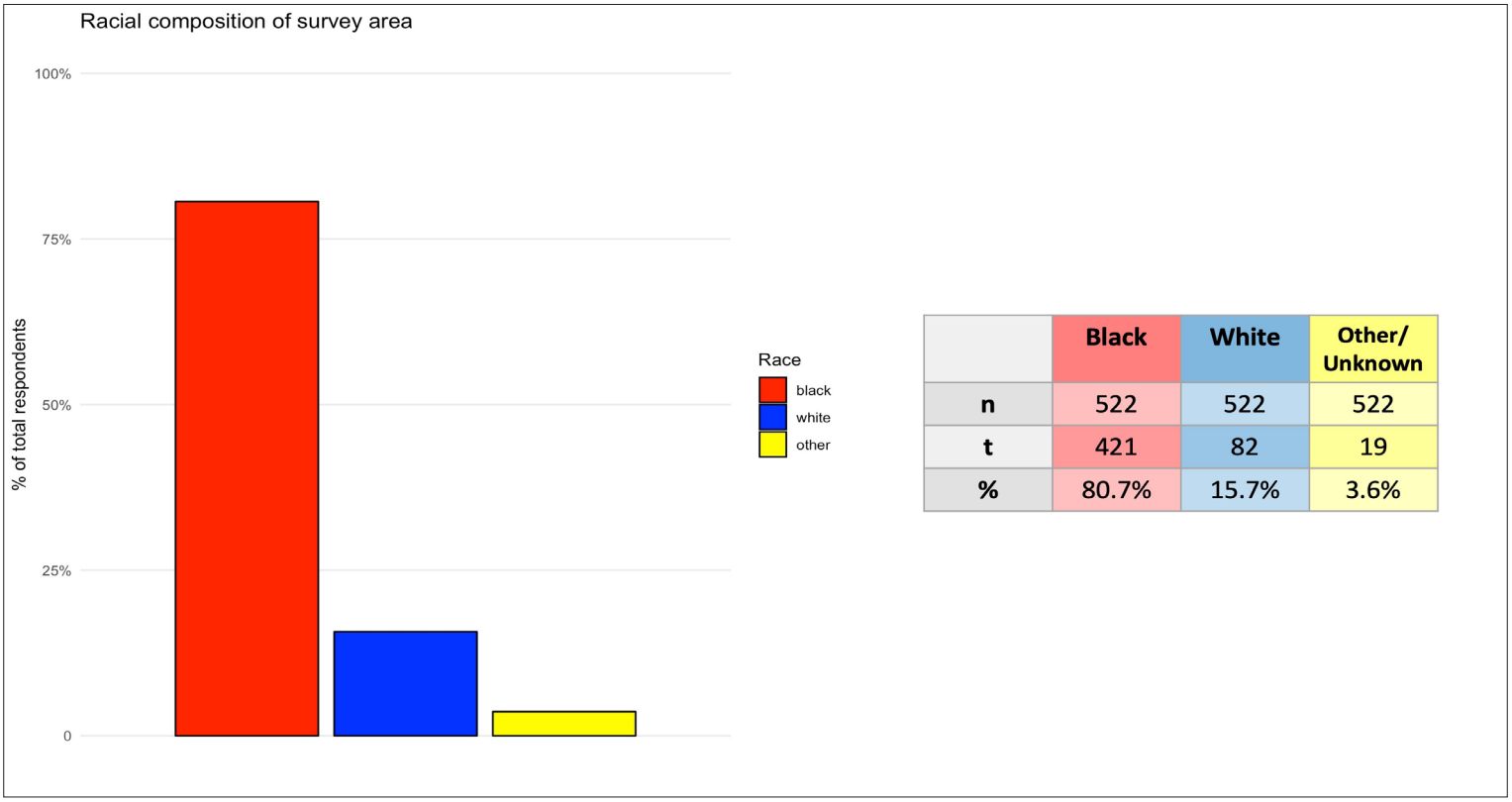
1 Survey implementers reported, for example, that respondents were sometimes unsure whether a household member had been "told by a doctor or another health care provider" that they had a certain condition (i.e. diagnosed with the condition), or simply whether the household member believed they had the condition. In such cases, survey implementers did *not* record a diagnosis.

Race / Ethnicity
Composition of Survey Area

Race / Ethnicity Composition of Survey Area

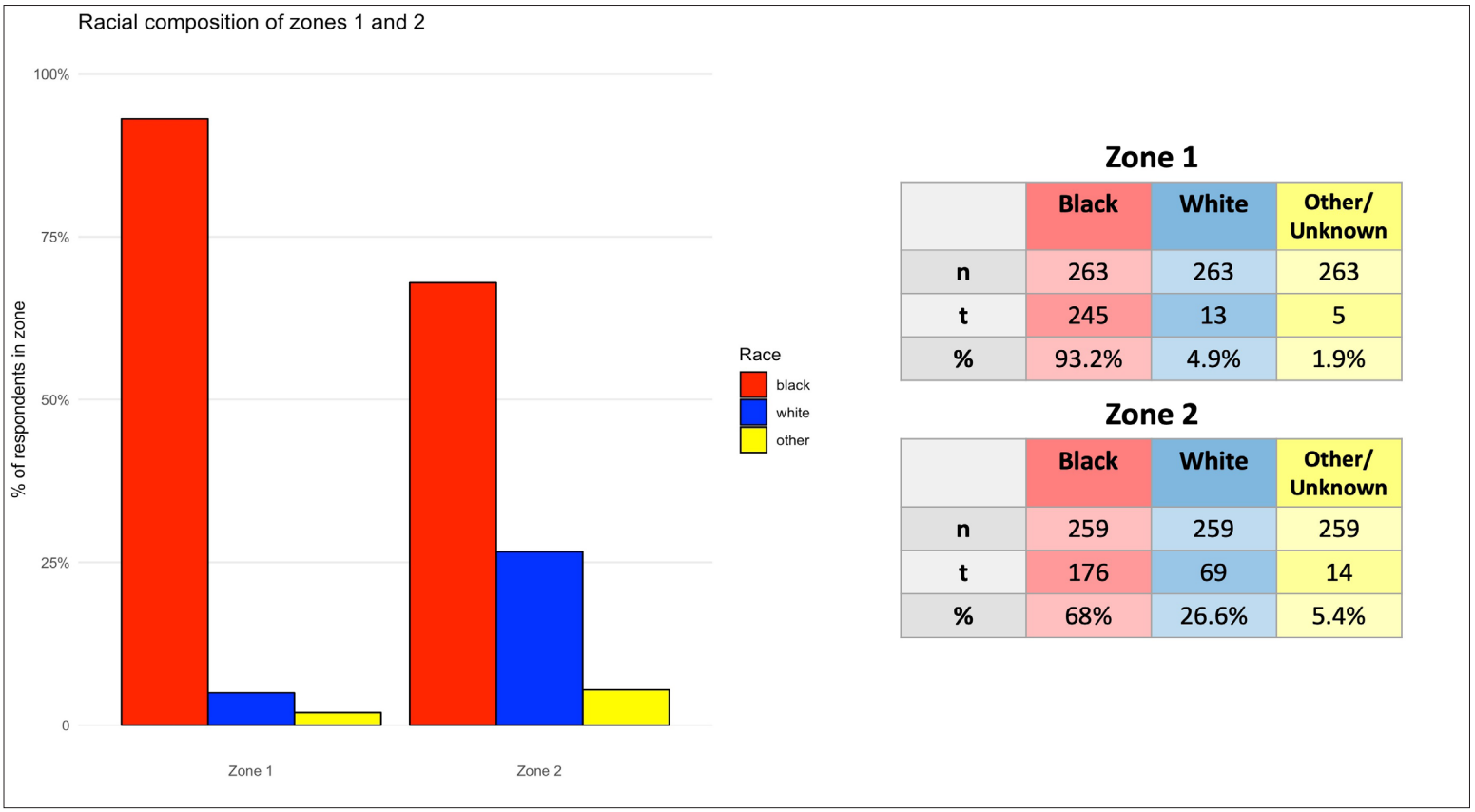
Self-reported race/ethnicity data were collected for respondents only. Respondents could select one or multiple of the following race/ethnicity categories on the survey: Asian; Black or African-American; Hispanic/Latina/Latino/Latinx; Native American; Native Hawaiian and Other Pacific Islander; White; Other.

The following visualizations provide a race/ethnicity breakdown of (1) full-time[1] respondents throughout the survey area, and (2) full-time respondents by spatial zone.



	Black	White	Other/ Unknown
n	522	522	522
t	421	82	19
%	80.7%	15.7%	3.6%

1 Respondents (and residents) who live in the household for 6-7 days of the week are considered "full-time."



Zone 1

	Black	White	Other/ Unknown
n	263	263	263
t	245	13	5
%	93.2%	4.9%	1.9%

Zone 2

	Black	White	Other/ Unknown
n	259	259	259
t	176	69	14
%	68%	26.6%	5.4%

Introduction to Monte Carlo Analyses of Cancer Prevalence

Introduction to Monte Carlo Analyses of Cancer Prevalence

DESCRIPTION

We used the Monte Carlo simulation method to analyze our health survey data on cancer prevalence among respondents, cancer prevalence among residents, clustering of cancer diagnoses within households, and childhood cancer prevalence.

A Monte Carlo simulation is a statistical method that models the probabilities of different outcomes so as to account for inherent uncertainty—that is, the intervention of random variables that may alter any potential outcome at any given time. We used a computer program to simulate a population in the United States with the same race, sex, and age demographics as the sample we surveyed. Using the Monte Carlo method, the computer program modeled the probabilities of different cancer rates we might plausibly see in the simulated population. This enabled us to compare the cancer rates our team documented with the cancer rates we are likely to see—based on national cancer statistics broken down by race, sex, and age—in a population in the United States that looks like the sample we surveyed.

METHODOLOGY

For every resident in our survey sample, we had a corresponding resident—of the same race, sex, and age—in our simulated population. Each member of the simulated population was assigned a value of 0 (no cancer diagnosis in the previous 23 years) or 1 (one or more cancer diagnoses in the previous 23 years). The probability that a simulated resident in a certain race/sex/age group would be assigned 0 or 1 was based on the National Cancer Institute's 2015 Surveillance, Epidemiology, and End Results (SEER) data (see *Source for National Cancer Data,* below).[1]

For example: According to SEER data, 23-year cancer prevalence among Black men between the ages of 60 and 69 is about 12.8%. In our simulated population, every Black male in his 60s was randomly assigned a value of 1 with probability p = 12.8% (otherwise, a value of 0 with probability 1-p = 87.2%).

1 This section focuses on our methodology for individual-level (respondent/resident) Monte Carlo analyses. We also conducted a Monte Carlo analysis of household-level clustering of cancer diagnoses (see *Clustering of cancer diagnoses within households*, below). In those analyses, for every *household* in our survey pool, we had a corresponding household—with household members of the same race, sex, and age—in our simulated pool. Each *household* was assigned a value of 0 or 1.

Each simulated resident was assigned a value of 0 or 1 in this manner, using the SEER cancer prevalence data for that resident's race/sex/age group. The process was then repeated 9,999 times to generate a total of 10,000 simulations. This enabled us to compare the cancer prevalence outcome of our survey sample to a *distribution* of cancer prevalence outcomes in our simulated population and determine the probability that the cancer prevalence outcome of our survey sample would occur in a simulated population with the same race, sex, and age makeup. In other words, we used SEER data to determine the likelihood of a range of cancer prevalence outcomes that we might see in our survey sample; then we compared the actual cancer prevalence in our survey sample to this range.

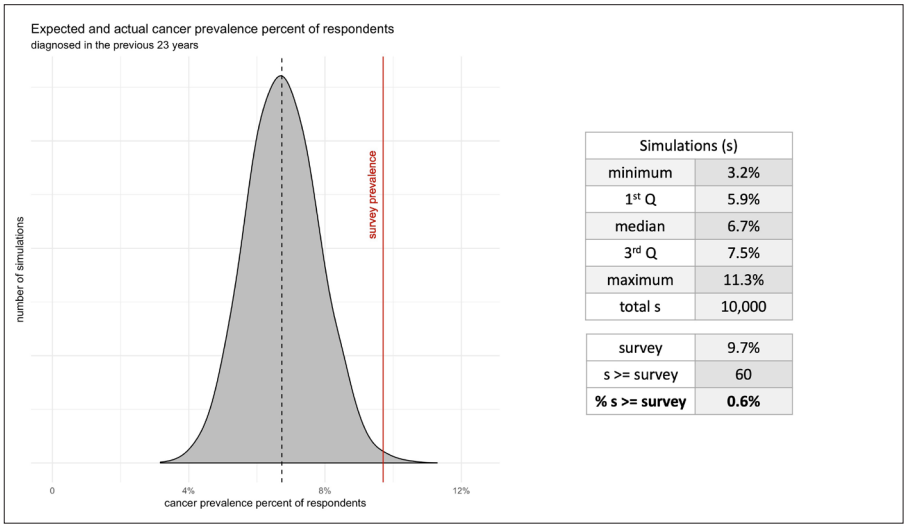
Race, sex, and age were considered in our Monte Carlo analyses because SEER data are broken down by these three demographic variables. Other demographic variables (such as socioeconomic status) could not be considered because we lacked comparable national-level cancer prevalence data for other variables.

HOW TO INTERPRET MONTE CARLO GRAPHS

Overview

When viewing the distribution of cancer prevalence outcomes in our simulated population (see **Sample Figure 1.1**), note that the *median* of the distribution (represented by the dotted vertical line in the center) is an approximation of the most likely cancer prevalence outcome. Half the simulations yielded cancer prevalence outcomes higher than the median and half the simulations yielded cancer prevalence outcomes lower than the median. The solid red vertical line (labeled "survey prevalence") represents the *actual* cancer prevalence in our survey sample. The greater the distance is between the solid red line (actual cancer prevalence) and the dotted line (approximation of most likely cancer prevalence), the *more unusual* the cancer prevalence in our survey sample. In Sample Figure 1.1, only 0.6% of the simulations yielded cancer prevalence outcomes higher than the cancer prevalence in our survey sample; that is, only 0.6% of the simulations yielded cancer prevalence outcomes to the right of the solid line. This means that the probability that a population with the same race, sex, and age composition as our survey sample would have a cancer prevalence greater than or equal to that of our survey sample is only 0.6%—extremely unlikely. This 0.6% probability is called the "p-value." **The lower the p-value, the more unusual the outcome.** P-values of 5% or less are considered statistically significant.

SAMPLE FIGURE 1.1



Simulations (s)	
minimum	3.2%
1st Q	5.9%
median	6.7%
3rd Q	7.5%
maximum	11.3%
total s	10,000

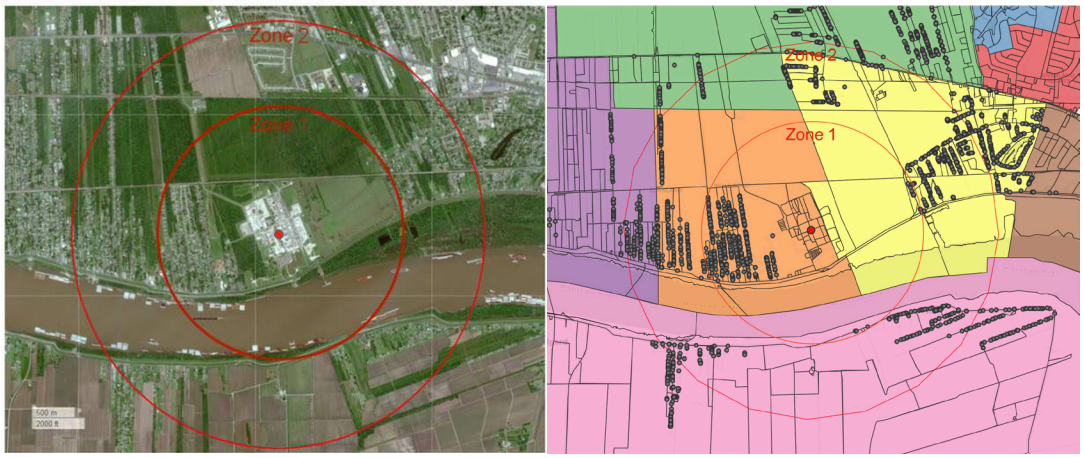
survey	9.7%
s >= survey	60
% s >= survey	**0.6%**

Each of our graphs is accompanied by a corresponding table of the following values (see **Sample Figure 1.1**):

- Minimum: lowest cancer prevalence in the simulation distribution
- First quartile: cancer prevalence value at which 25% of the simulations yielded lower values and 75% of the simulation yielded higher values
- Median: cancer prevalence value at which 50% of the simulations yielded lower values and 50% of the simulations yielded higher values
- Third quartile: cancer prevalence value at which 75% of the simulations yielded lower values and 25% of the simulations yielded higher values
- Maximum: highest cancer prevalence value in the simulation distribution
- Total s: total number of simulations (always 10,000)
- Survey: cancer prevalence in the survey sample
- s >= survey: number of simulation values greater than or equal to the survey sample cancer prevalence
- **% s >= survey: percentage of simulation values greater than or equal to the survey sample cancer prevalence (i.e. the "p-value," or the probability that a simulated population with the same demographic makeup as our survey sample would have a cancer prevalence greater than or equal to that of our survey sample)**

Spatial Zone Analysis

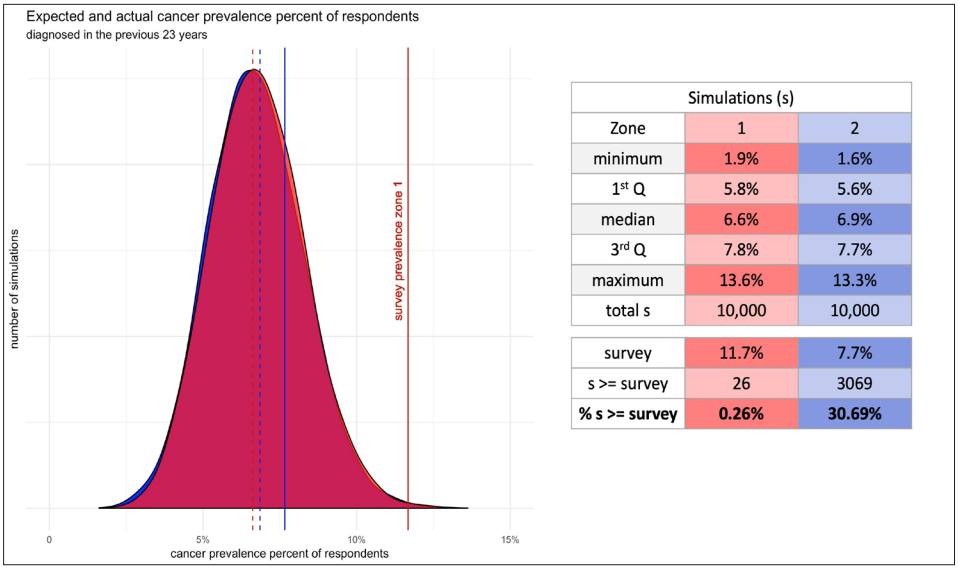
As noted above, we surveyed 60% of the households within a 1.5-kilometer radius of the Denka facility ("Zone 1") and 20% of the households located between 1.5 and 2.5 kilometers from the facility ("Zone 2").



Maps of the survey area. The outer circle circumscribes the area within a 2.5-km radius of the Denka facility. The inner circle circumscribes the area within a 1.5-km radius of the Denka facility. The facility can be seen at the center of the survey area. In the map on the right, the grey dots represent households and each color represents a different census tract. Residents of the orange-colored census tract face the highest risk in the country of developing cancer from air pollution, according to the EPA.

In addition to using Monte Carlo simulations to analyze overall cancer prevalence in the survey area, we conducted analyses by spatial zone to determine the relationship, if any, between cancer prevalence and proximity to the Denka facility. These analyses enabled us to separately determine—and compare—the cancer prevalence probabilities closer to the plant (in Zone 1) and farther away from the plant (in Zone 2).[1]

When viewing spatial zone distributions of cancer prevalence values in our simulated population (see **Sample Figure 1.2**), note that each zone is represented by its own color-coded distribution: the red distribution shows the range of cancer prevalence values we might see in Zone 1, and the blue distribution shows the range of cancer prevalence values we might see in Zone 2. The dotted red line represents the median of the Zone 1 distribution, and the dotted blue line represents the median of the Zone 2 distribution. The solid red line (labeled "survey prevalence zone 1") represents the cancer prevalence of the Zone 1 survey sample, and the solid blue line represents the cancer prevalence of the Zone 2 survey sample.



Simulations (s)		
Zone	1	2
minimum	1.9%	1.6%
1st Q	5.8%	5.6%
median	6.6%	6.9%
3rd Q	7.8%	7.7%
maximum	13.6%	13.3%
total s	10,000	10,000
survey	11.7%	7.7%
s >= survey	26	3069
% s >= survey	**0.26%**	**30.69%**

SAMPLE FIGURE 1.2

Because there is not a significant difference in the range of expected cancer prevalence outcomes for Zone 1 and Zone 2, the two distributions overlap significantly, and their medians are always clustered together (and sometimes even exactly aligned).

1 We also analyzed cancer prevalence probabilities in the area within 1.25 kilometers of the plant ("Zone A") and the area between 1.25 and 2.5 kilometers from the plant ("Zone B"). See Appendix for Zone A and Zone B Monte Carlo analyses.

In Sample Figure 1.2, the Zone 1 survey prevalence line is much further from its corresponding median than is the Zone 2 survey prevalence line. This means that the cancer prevalence for Zone 1—the zone closer to the plant—is significantly more unusual/improbable than the cancer prevalence for Zone 2. This can be confirmed using the graph's corresponding table of zone-specific values (see **Sample Figure 1.2**), which lists the Zone 1 p-value (i.e. probability of the Zone 1 cancer prevalence outcome) at 0.26%—extremely unlikely—and the Zone 2 p-value (i.e. probability of the Zone 2 cancer prevalence outcome) at 30.69%.

Smoking Exclusion Criterion

For our respondent and resident Monte Carlo analyses, we used a smoking exclusion criterion. This exclusion criterion removed all respondents/residents who live in households where anyone smokes on a daily basis. Since corresponding respondents/residents were also removed from the simulated population, the smoking exclusion criterion impacted the range of simulated outcomes as well as the survey outcome.

DATA SET USED FOR MONTE CARLO ANALYSES

All part-time respondents and residents (defined as those who live in the household for only 1-5 days of the week, inclusive) were eliminated from the data set. The one household that consisted exclusively of part-time residents was also eliminated.

Although race information was collected for respondents only, we assumed—for purposes of the Monte Carlo analyses only—that all residents of a household shared the race of the respondent.

Any respondents and residents for whom we did not have all three pieces of necessary demographic information—race, sex, and age—were eliminated from the data set. In addition, 21 respondents and residents who reported a race/ethnicity for which there is no SEER analogue (and therefore no comparative national cancer prevalence statistic) were eliminated from the data set. Finally, since we used SEER's 23-year cancer prevalence statistics, we eliminated respondents and residents whose only cancer diagnoses happened in 1994 or earlier (more than 23 years prior to the health survey).

As a result of this process, if a particular respondent was eliminated from the data set (either because we had no race information/non-comparable race information for the respondent, or because the respondent's only cancer diagnoses occurred more than 23 years ago), all members of the respondent's household were eliminated from the data set as well (since the other household members' race depended on the respondent's race). Since all members of the respondent's household were eliminated from the data set, the household itself was eliminated. 12 households were eliminated from the data set in this manner.

After all eliminations, the total numbers of respondents, residents, and households included in the Monte Carlo analyses were as follows:

	Zone 1	Zone 2	Total
Households	262	263	525
Respondents	257	248	505
Residents	777	730	1,507

Race/ethnicity selections on the survey instrument were paired with SEER counterparts as follows (numbers of survey residents in each race/ethnicity category are shown in parentheses):

Survey	SEER
Black or African-American [alone or with any other category] **(1292)**	Black
White [alone or with any other category except Black or African-American] **(190)**	White
Hispanic/Latina/Latino/Latinx [alone or with any other category except Black or African American or White] **(25)**	Hispanic
Asian OR Native Hawaiian and other Pacific Islander [alone or with any other category except Black or African-American, White, or Hispanic/Latina/Latino/Latinx] **(0)**	Asian/Pacific Islander

SOURCE FOR NATIONAL CANCER DATA

Noone AM, Howlader N, Krapcho M, Miller D, Brest A, Yu M, Ruhl J, Tatalovich Z, Mariotto A, Lewis DR, Chen HS, Feuer EJ, Cronin KA (eds). SEER Cancer Statistics Review, 1975-2015, National Cancer Institute. Bethesda, MD, https://seer.cancer.gov/csr/1975_2015/, based on November 2017 SEER data submission, posted to the SEER web site, April 2018.

Results of Monte Carlo Analyses

Results of Monte Carlo Analyses

The results of our Monte Carlo analyses are presented in the following sections:

(I) Cancer Prevalence Among Respondents:
One member of each household surveyed (the "respondent") was asked for health and demographic information about themself and all other members of the household. This section presents our analyses of cancer prevalence among *respondents only*.

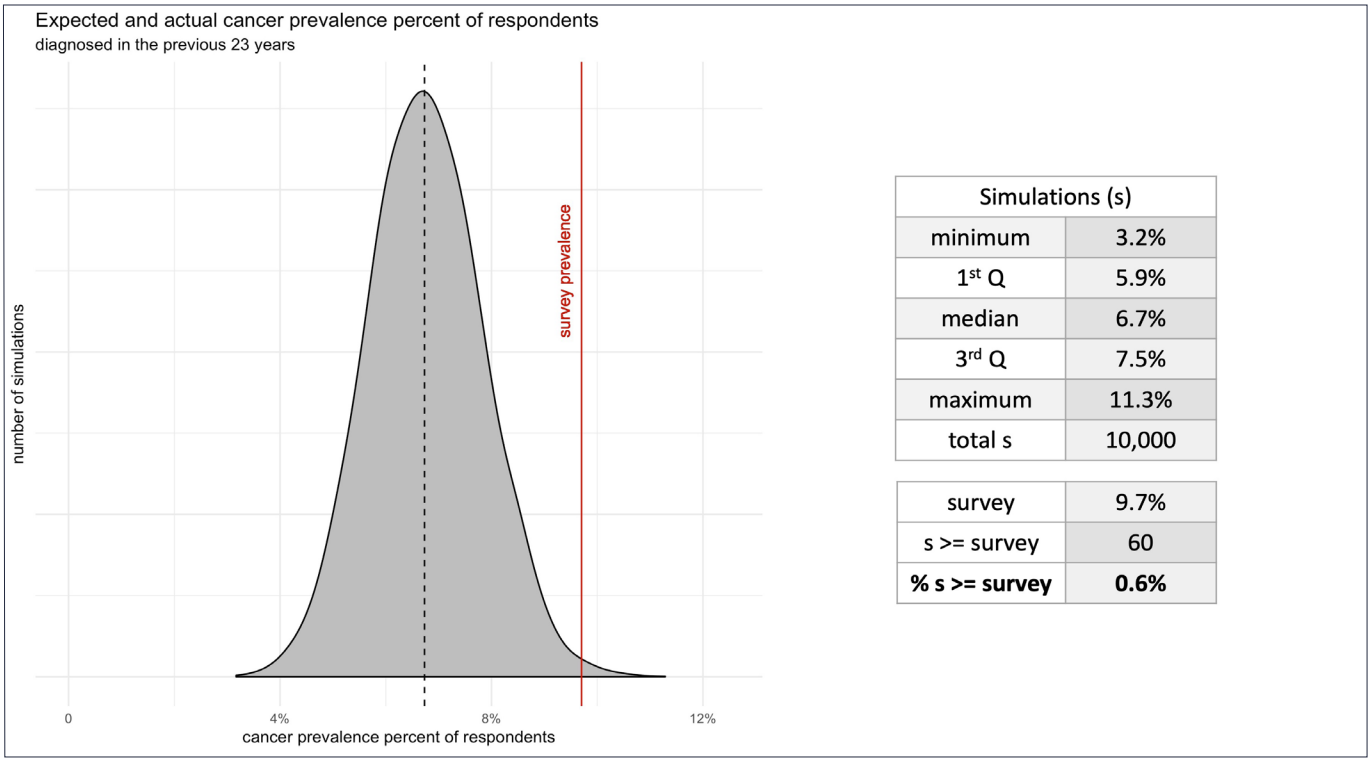
(II) Cancer Prevalence Among Residents:
This section presents our analyses of cancer prevalence among "residents" (respondents plus their household members)—that is, all individuals about whom information was collected.

(III) Clustering of Cancer Diagnoses Within Households:
This section presents our analysis of the percent of households with at least two residents who have been diagnosed with cancer.

(IV) Childhood Cancer Prevalence:
This section presents our analysis of childhood cancer prevalence.

(I) CANCER PREVALENCE AMONG RESPONDENTS, without and with smoking exclusion criterion

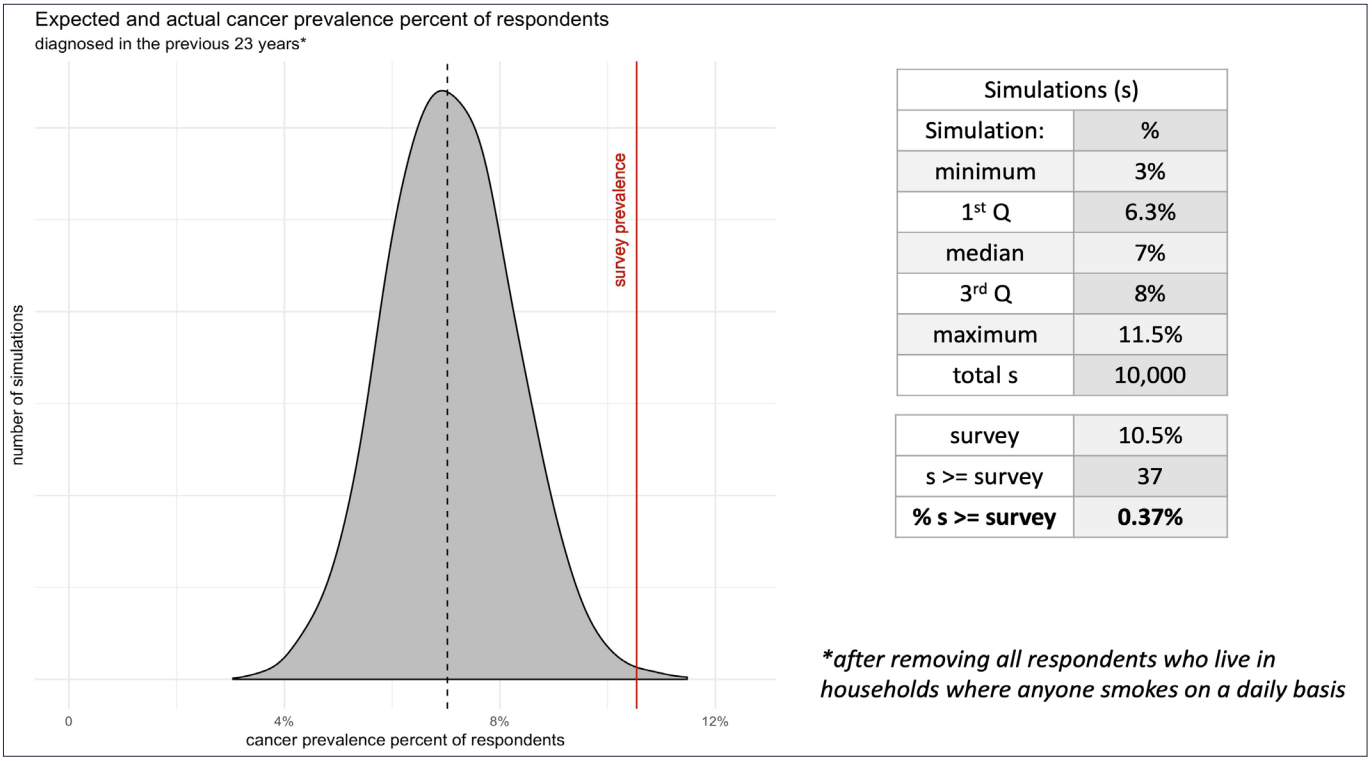
The graphs below show expected cancer prevalence among members of a simulated population with the same race, sex, and age demographics as our survey sample of respondents. The graphs also show actual cancer prevalence among survey respondents, and the likelihood of this prevalence.



Simulations (s)	
minimum	3.2%
1st Q	5.9%
median	6.7%
3rd Q	7.5%
maximum	11.3%
total s	10,000

survey	9.7%
s >= survey	60
% s >= survey	**0.6%**

FIGURE 2.1 (ABOVE): Very statistically significant p-value, indicating that 9.7% cancer prevalence among respondents is not due to chance

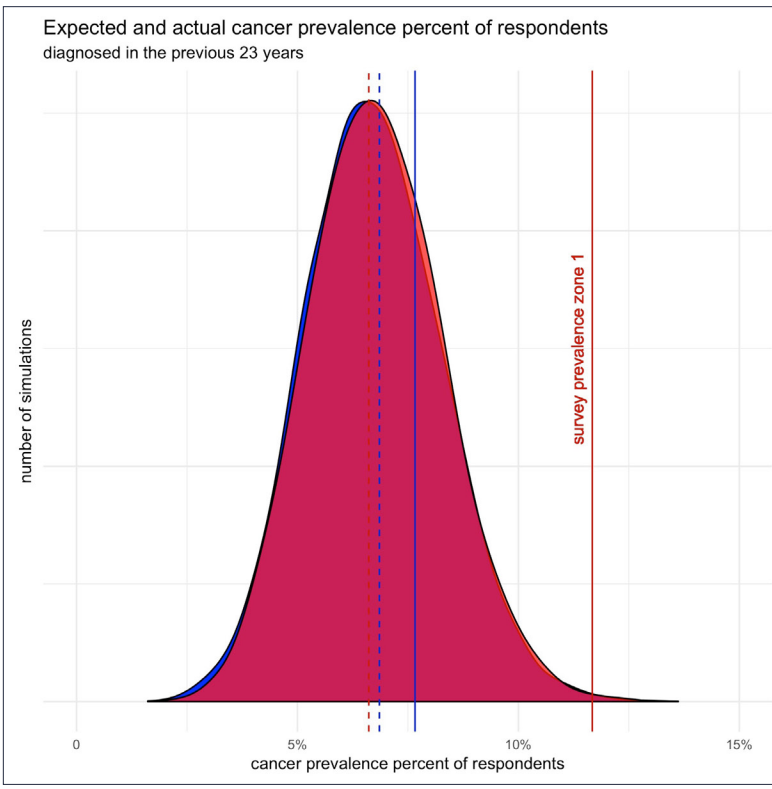


Simulations (s)	
Simulation:	%
minimum	3%
1st Q	6.3%
median	7%
3rd Q	8%
maximum	11.5%
total s	10,000

survey	10.5%
s >= survey	37
% s >= survey	**0.37%**

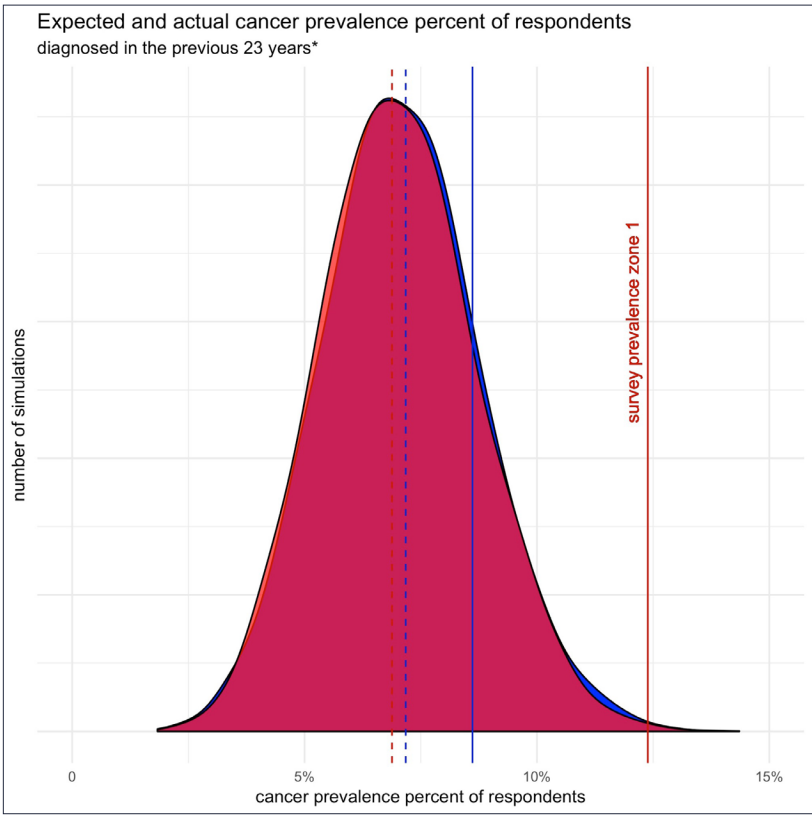
**after removing all respondents who live in households where anyone smokes on a daily basis*

FIGURE 2.2 (ABOVE): Very statistically significant p-value (with smoking exclusion criterion), indicating that 10.5% cancer prevalence among respondents is not due to chance



Simulations (s)		
Zone	1	2
minimum	1.9%	1.6%
1st Q	5.8%	5.6%
median	6.6%	6.9%
3rd Q	7.8%	7.7%
maximum	13.6%	13.3%
total s	10,000	10,000
survey	11.7%	7.7%
s >= survey	26	3069
% s >= survey	**0.26%**	**30.69%**

FIGURE 2.3 (ABOVE): Very statistically significant p-value for Zone 1, indicating that 11.7% cancer prevalence among Zone 1 respondents is not due to chance



Simulations (s)		
Zone	1	2
minimum	1.8%	1.9%
1st Q	6%	5.7%
median	6.9%	7.2%
3rd Q	8.3%	8.1%
maximum	14.2%	14.4%
total s	10,000	10,000
survey	12.4%	8.6%
s >= survey	22	2165
% s >= survey	**0.22%**	**21.65%**

**after removing all respondents who live in households where anyone smokes on a daily basis*

FIGURE 2.4 (ABOVE): Very statistically significant p-value for Zone 1 (with smoking exclusion criterion), indicating that 12.4% cancer prevalence among Zone 1 respondents is not due to chance

Figures 2.1 and 2.2 show respondent data for the survey area as a whole, while **Figures 2.3 and 2.4** show respondent data by spatial zone.

In **Figure 2.1**, the median of the curve (depicted by the dotted vertical line) is 6.7%, which means that there is a 50% chance that cancer prevalence in a simulated population with the same race, sex, and age makeup as our respondent group would be greater than 6.7%. As the corresponding table of values shows, the third quartile cancer prevalence is 7.5%; that is, there is only a 25% chance that cancer prevalence would be greater than 7.5%.

The actual cancer prevalence in our respondent group (depicted by the solid vertical line) is 9.7%. **The odds that cancer prevalence would be this high in a population with the same race, sex, and age demographics (the "p-value") is only 0.6%—a little over one-half of one percent.** In other words, the cancer prevalence among survey respondents is highly unusual—there is only a 0.6% likelihood that a cancer prevalence value this high could be due to chance. This is far below the 5% threshold for statistical significance.

Figure 2.2 shows respondent data for the survey area as a whole *after removal of respondents who live in households where anyone smokes on a daily basis*. Cancer prevalence among respondents actually *increases* to 10.5% when those exposed to smoke on a daily basis are removed from the pool. The p-value (i.e. odds that cancer prevalence would be this high in a population with the same race, sex, and age demographics) drops to 0.37%—a little over one-third of one percent. This is far below the 5% threshold for statistical significance.

In **Figure 2.3**, there are two distributions—one for Zone 1 (red) and another for Zone 2 (blue). Cancer prevalence among respondents in Zone 1 is 11.7%. The likelihood that cancer prevalence would be this high (i.e. the p-value) is only 0.26%. In other words, a cancer prevalence of 11.7% in Zone 1 is almost certainly *not* due to chance. A 0.26% p-value is far below the 5% threshold for statistical significance.

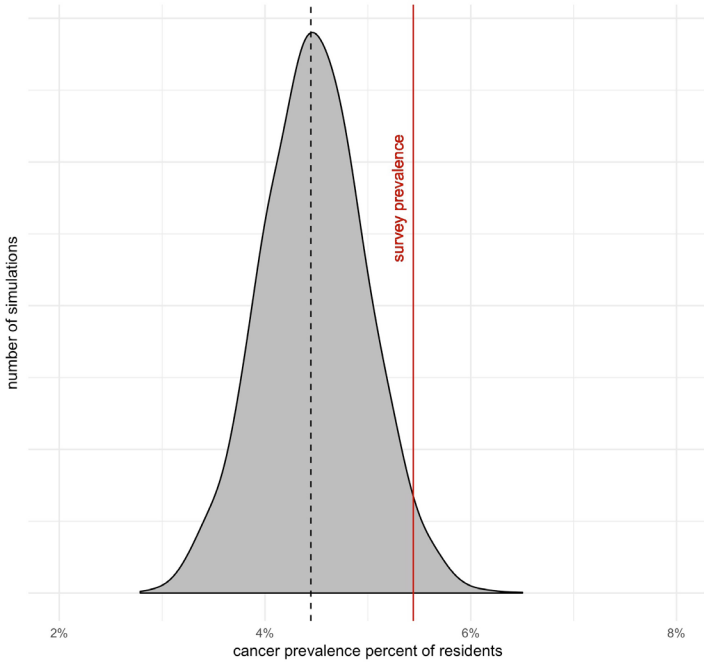
In Zone 2, cancer prevalence drops to 7.7%. Although this cancer prevalence value is still high—it is equivalent to the zone's third quartile simulation value—**the drop in cancer prevalence (and increase in p-value) from Zone 1 to Zone 2 indicates that there is an association between closer proximity to the Denka facility and higher cancer prevalence among survey respondents.**

Figure 2.4 shows respondent data by spatial zone *after removal of respondents who live in households where anyone smokes on a daily basis*. Cancer prevalence among respondents in Zone 1 actually *increases* to 12.4% when those exposed to smoke on a daily basis are removed from the pool. The p-value drops to 0.22%. This is far below the 5% threshold for statistical significance. Cancer prevalence among respondents in Zone 2 also increases when people exposed to daily smoke are removed.

(II) CANCER PREVALENCE AMONG RESIDENTS, without and with smoking exclusion criterion

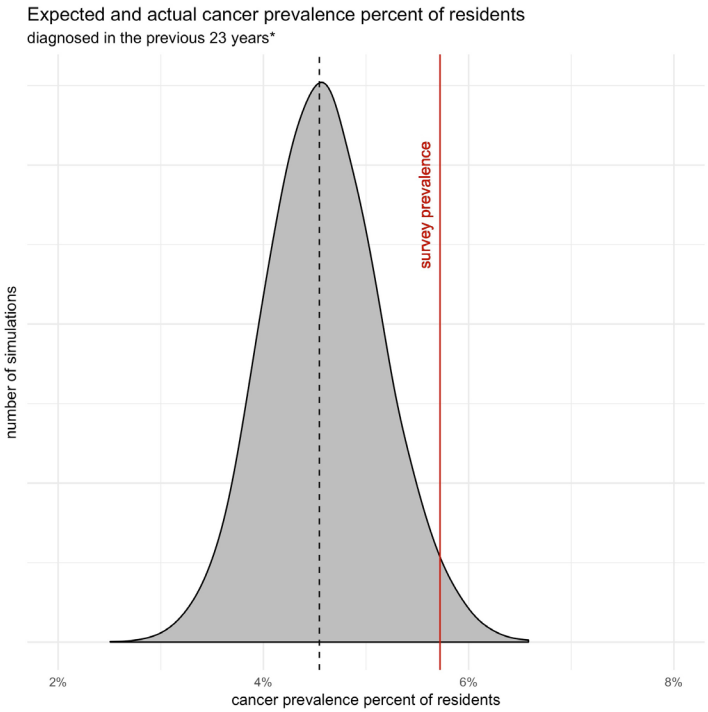
The graphs below show expected cancer prevalence among members of a simulated population with the same race, sex, and age demographics as our survey sample of residents (i.e. everyone for whom data were collected). The graphs also show actual cancer prevalence among residents surveyed, and the likelihood of this prevalence.

FIGURE 3.1 (RIGHT): Statistically significant p-value, indicating that 5.4% cancer prevalence among residents is not due to chance



Simulations (s)	
minimum	2.8%
1st Q	4.1%
median	4.4%
3rd Q	4.8%
maximum	6.5%
total s	10,000

survey	5.4%
s >= survey	343
% s >= survey	**3.43%**

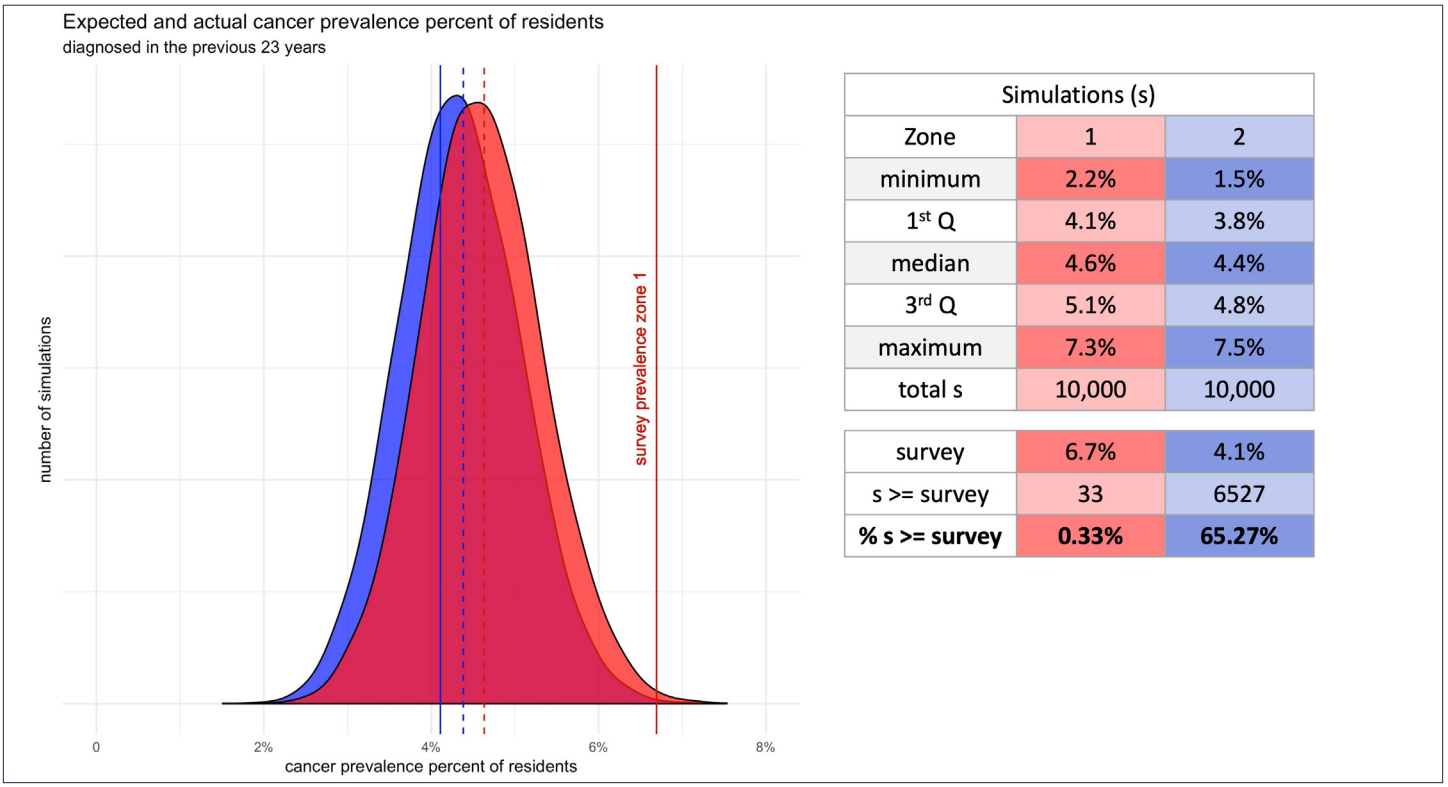


Simulations (s)	
minimum	2.5%
1st Q	4.2%
median	4.5%
3rd Q	4.9%
maximum	6.6%
total s	10,000

survey	**5.4%**
s >= survey	306
% s >= survey	**3.06%**

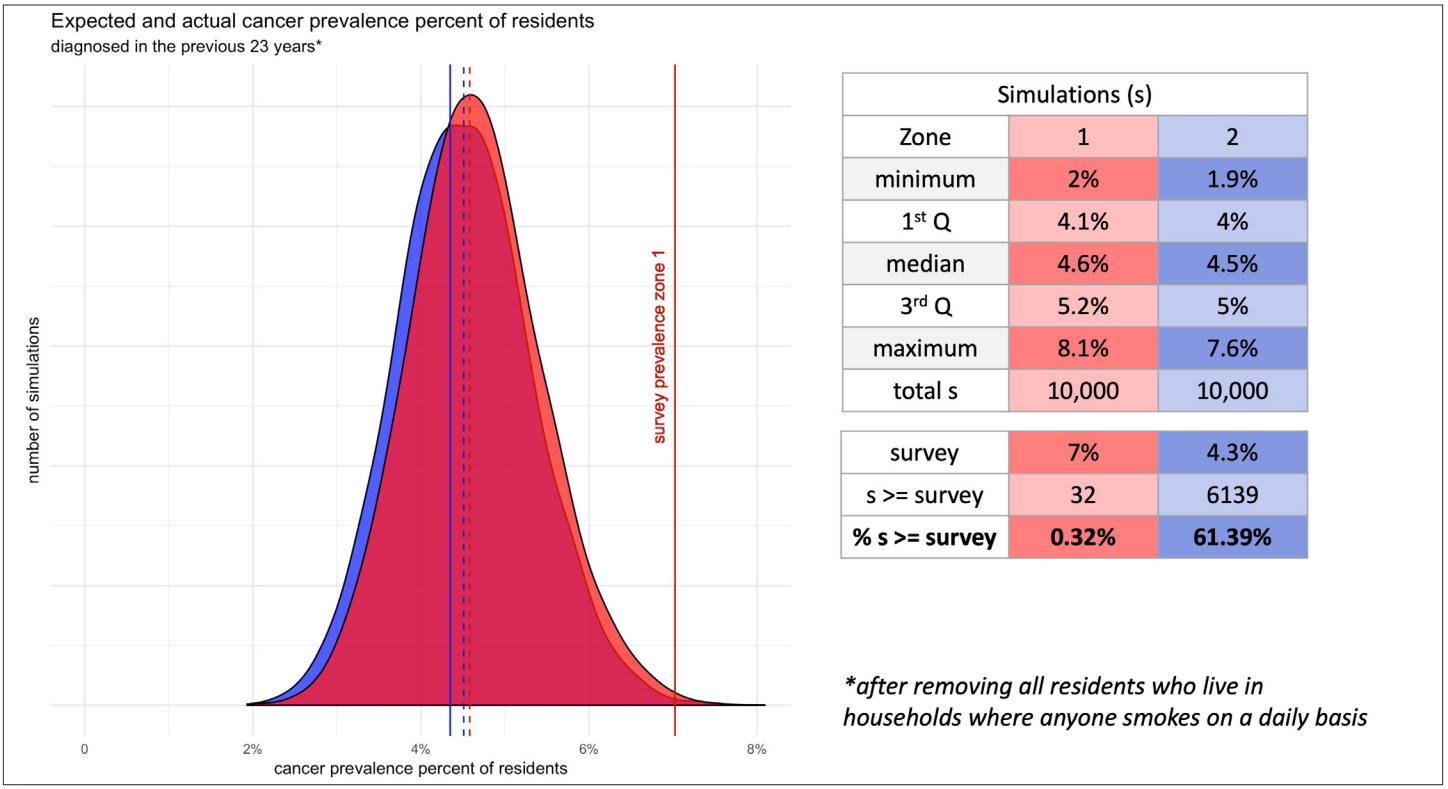
**after removing all residents who live in households where anyone smokes on a daily basis*

FIGURE 3.2 (LEFT): Statistically significant p-value (with smoking exclusion criterion), indicating that 5.4% cancer prevalence among residents is not due to chance



Simulations (s)		
Zone	1	2
minimum	2.2%	1.5%
1st Q	4.1%	3.8%
median	4.6%	4.4%
3rd Q	5.1%	4.8%
maximum	7.3%	7.5%
total s	10,000	10,000
survey	6.7%	4.1%
s >= survey	33	6527
% s >= survey	**0.33%**	**65.27%**

FIGURE 3.3 (ABOVE): Very statistically significant p-value for Zone 1, indicating that 6.7% cancer prevalence among Zone 1 residents is not due to chance



Simulations (s)		
Zone	1	2
minimum	2%	1.9%
1st Q	4.1%	4%
median	4.6%	4.5%
3rd Q	5.2%	5%
maximum	8.1%	7.6%
total s	10,000	10,000
survey	7%	4.3%
s >= survey	32	6139
% s >= survey	**0.32%**	**61.39%**

**after removing all residents who live in households where anyone smokes on a daily basis*

FIGURE 3.4 (ABOVE): Very statistically significant p-value for Zone 1 (with smoking exclusion criterion), indicating that 7% cancer prevalence among Zone 1 residents is not due to chance

Figures 3.1 and 3.2 show resident data for the survey area as a whole, while **Figures 3.3 and 3.4** show resident data by spatial zone.

In **Figure 3.1**, the median of the curve (depicted by the dotted vertical line) is 4.4%, which means that there is a 50% chance that cancer prevalence in a simulated population with the same race, sex, and age makeup as our survey sample would be greater than 4.4%. As the corresponding table of values shows, the third quartile cancer prevalence is 4.8%; that is, there is only a 25% chance that cancer prevalence would be greater than 4.8%.

The actual cancer prevalence in our survey sample (depicted by the solid vertical line) is 5.4%. **The odds that cancer prevalence would be this high in a population with the same race, sex, and age demographics (the "p-value") is only 3.43%.** In other words, the cancer prevalence among residents surveyed is highly unusual—there is only a 3.43% likelihood that a cancer prevalence value this high could be due to chance. This is below the 5% threshold for statistical significance.

Figure 3.2 shows cancer prevalence data for the survey area as a whole *after removal of residents who live in households where anyone smokes on a daily basis.* Cancer prevalence among residents surveyed remains the same when those exposed to smoke on a daily basis are removed from the pool. The p-value (i.e. odds that cancer prevalence would be this high in a population with the same race, sex, and age demographics) drops slightly, to 3.06%. This is below the 5% threshold for statistical significance.

In **Figure 3.3**, there are two distributions—one for Zone 1 (red) and another for Zone 2 (blue). Cancer prevalence among residents surveyed in Zone 1 is 6.7%. The likelihood that cancer prevalence would be this high (i.e. the p-value) is only 0.33%. In other words, a cancer prevalence of 6.7% in Zone 1 is almost certainly *not* due to chance. A 0.33% p-value is far below the 5% threshold for statistical significance.

In Zone 2, cancer prevalence drops to 4.1%. **The drop in cancer prevalence (and increase in p-value) from Zone 1 to Zone 2 indicates that there is an association between closer proximity to the Denka facility and higher cancer prevalence among residents surveyed.**

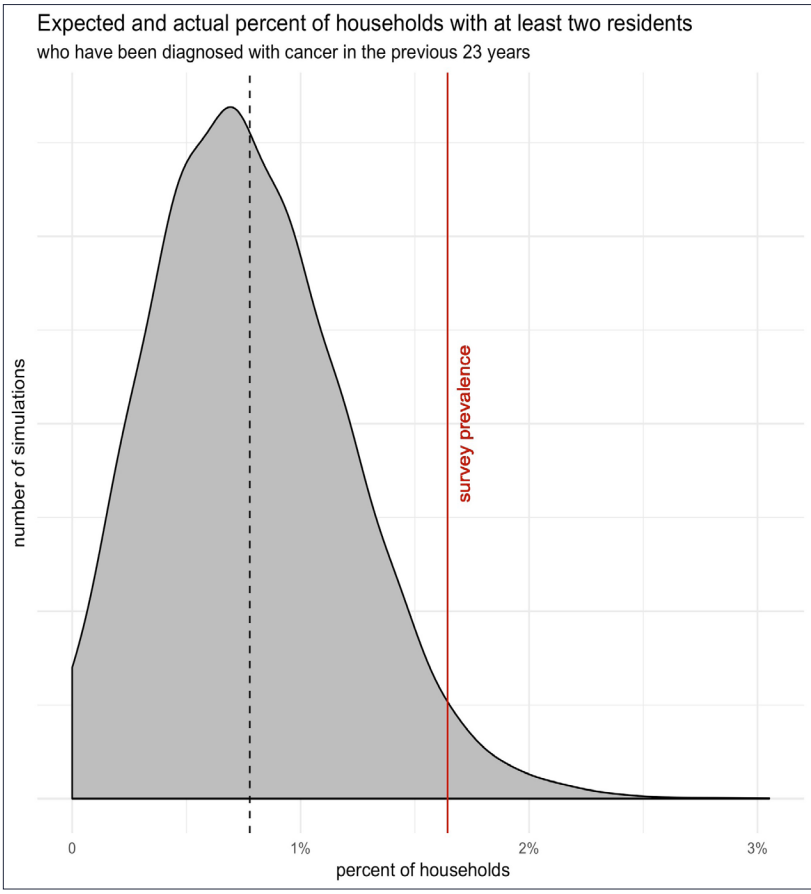
Figure 3.4 shows cancer prevalence data by spatial zone *after removal of residents who live in households where anyone smokes on a daily basis*. Cancer prevalence among residents surveyed in Zone 1 actually *increases* to 7% when those exposed to smoke on a daily basis are removed from the pool. The p-value drops slightly, to 0.32%. This is far below the 5% threshold for statistical significance. Cancer prevalence among residents surveyed in Zone 2 also increases when people exposed to daily smoke are removed.

(III) CLUSTERING OF CANCER DIAGNOSES WITHIN HOUSEHOLDS

This analysis uses Monte Carlo simulations at the household, rather than individual (respondent/resident), level. For every household in our survey pool, we had a corresponding household—with household members of the same race, sex, and age—in our simulated pool.

Figure 4.1 shows the expected proportion of households in this simulated pool with at least two residents who have been diagnosed with cancer. The figure also shows the *actual* proportion of households surveyed with at least two residents who have been diagnosed with cancer, as well as the likelihood of this proportion.

1.6% of households surveyed have at least two residents who have been diagnosed with cancer. The likelihood that this proportion of households would have two or more residents with cancer diagnoses (i.e. the likelihood that this was due to chance) is only 4.42%. This is below the 5% threshold for statistical significance. **This finding suggests an unusual clustering of cancer diagnoses within the same households. In an unusually high proportion of households, multiple members of the same household have been diagnosed with cancer.**



Simulations (s)	
minimum	0%
1st Q	0.5%
median	0.7%
3rd Q	0.9%
maximum	3.1%
total s	10,000

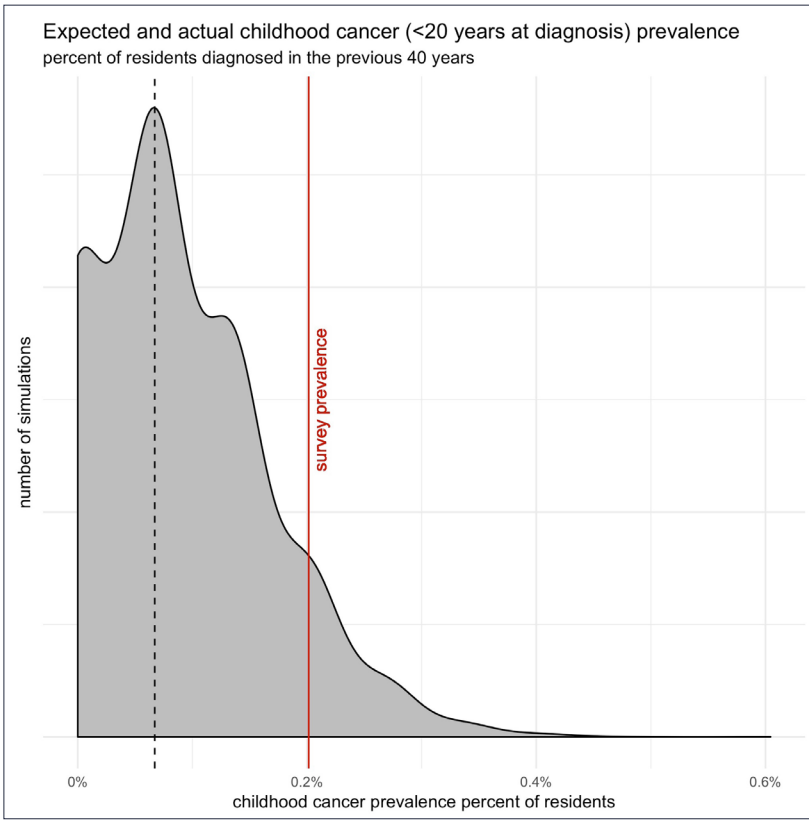
survey	1.6%
s >= survey	442
% s >= survey	**4.42%**

FIGURE 4.1 (ABOVE): Statistically significant p-value, indicating that clustering of cancer diagnoses within households (i.e. two or more members of the same household with cancer diagnoses) is not due to chance

(IV) CHILDHOOD CANCER PREVALENCE

Figure 5.1 shows expected childhood cancer prevalence among members of a simulated population with the same race and sex demographics as our survey sample.[1] The figure also shows actual childhood cancer prevalence among residents surveyed (counting residents who were diagnosed with cancer in the past 40 years and younger than 20 years old at the time of diagnosis), and the likelihood of this prevalence.

As the graph and corresponding table show, childhood cancer prevalence among residents surveyed is 0.2%, higher than the third quartile simulation value of 0.13%. The p-value (likelihood of a 0.2% childhood cancer prevalence) is only 13.93%, suggesting that childhood cancer prevalence in our survey sample may be unusually high.



Simulations (s)	
minimum	0%
1st Q	0%
median	0.07%
3rd Q	0.13%
maximum	0.6%
total s	10,000

survey	0.2%
s >= survey	1393
% s >= survey	**13.93%**

FIGURE 5.1 (ABOVE): Although not statsitically significant, p-value indicates that childhood cancer prevalence may be unusually high

1 A few notes of explanation regarding this analysis: (1) Black and white are the only races for which SEER statistics for childhood cancer prevalence are available. As a result, only Black males, Black females, white males, and white females were included in this analysis. All other members of our survey sample were eliminated from the pool. (2) SEER age statistics are not available for childhood cancer prevalence, so age was not considered. Since age was not considered, residents for whom we did not have age information were not eliminated from the pool; only residents for whom we did not have race or sex information were eliminated. (3) After eliminations, there were a total of 1,489 residents included in the Monte Carlo analysis for childhood cancer prevalence.

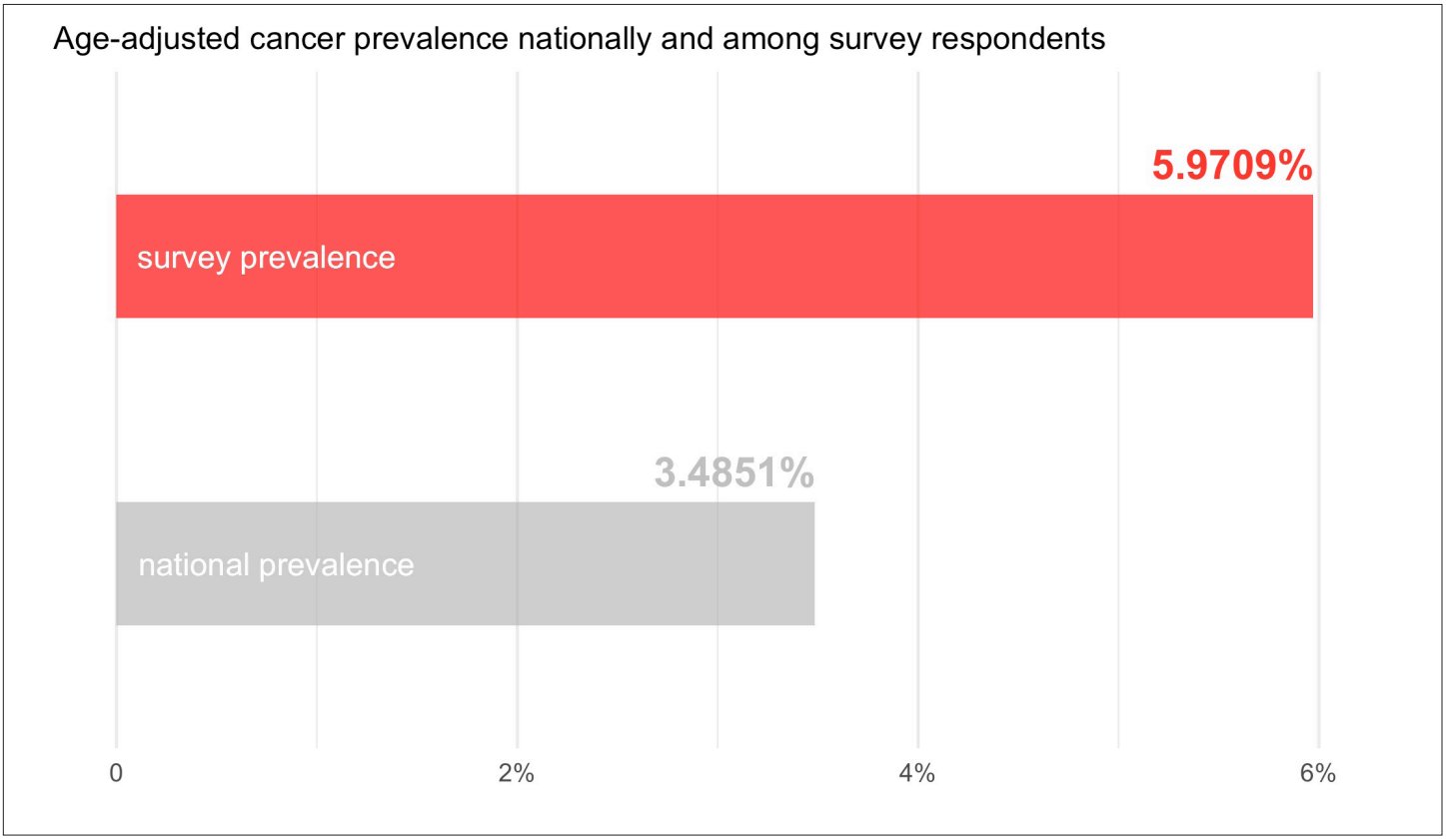
Comparison of Respondent / Resident and National Cancer Prevalence

Comparison of Respondent / Resident and National Cancer Prevalence

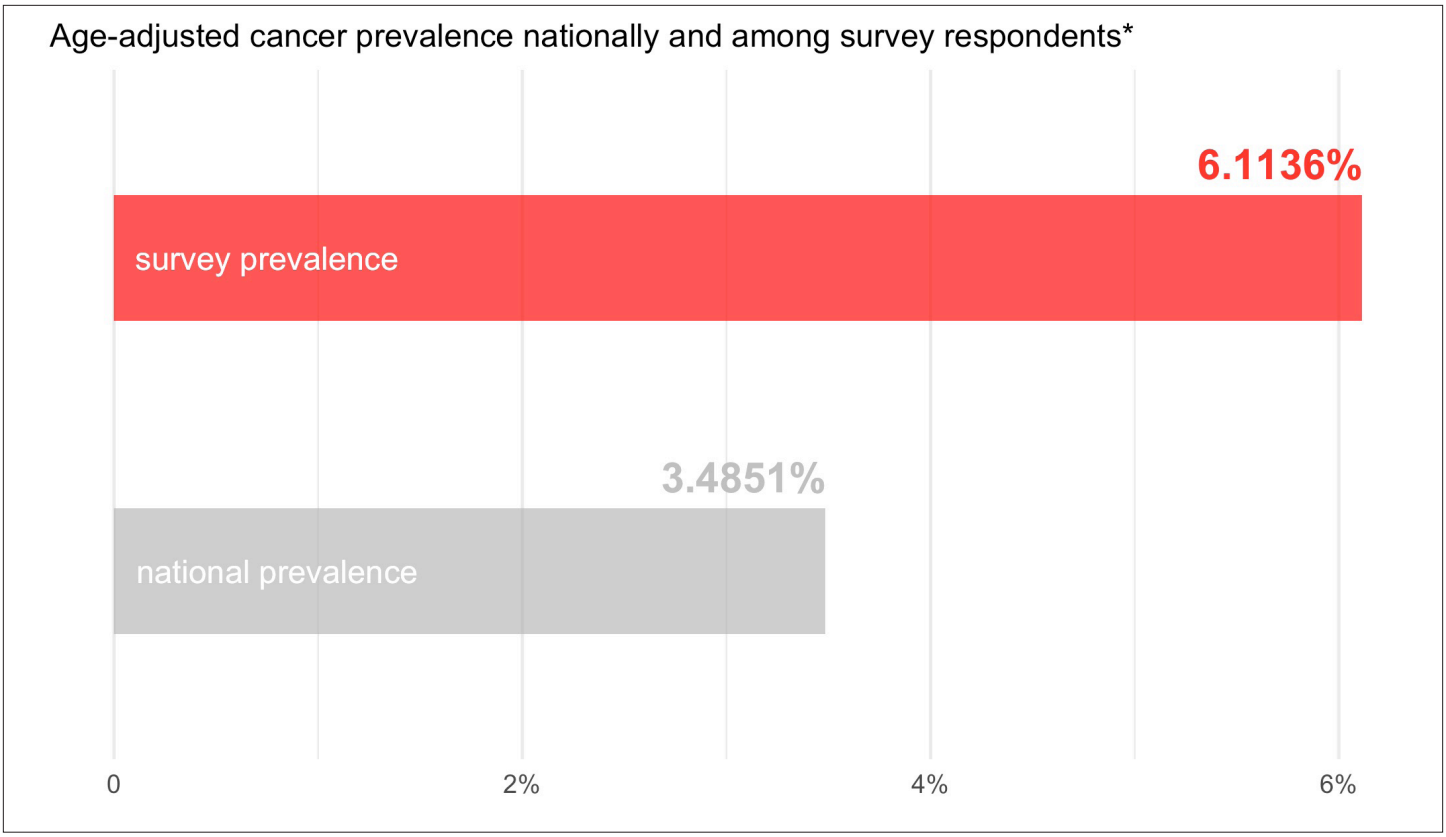
In this section, we present comparisons of our age-adjusted cancer prevalence data for survey respondents and residents with the SEER age-adjusted national cancer prevalence rate of 3.4851% (SEER Cancer Statistics Review, 1975-2015).

Crude cancer prevalence rates for survey respondents and residents were age-adjusted to the US Standard Population in the year 2000. This enabled comparison of the age-adjusted survey rates with the national rate, also age-adjusted (by SEER) to the 2000 US Standard Population.

(I) RESPONDENT VERSUS NATIONAL CANCER PREVALENCE, without and with smoking exclusion criterion



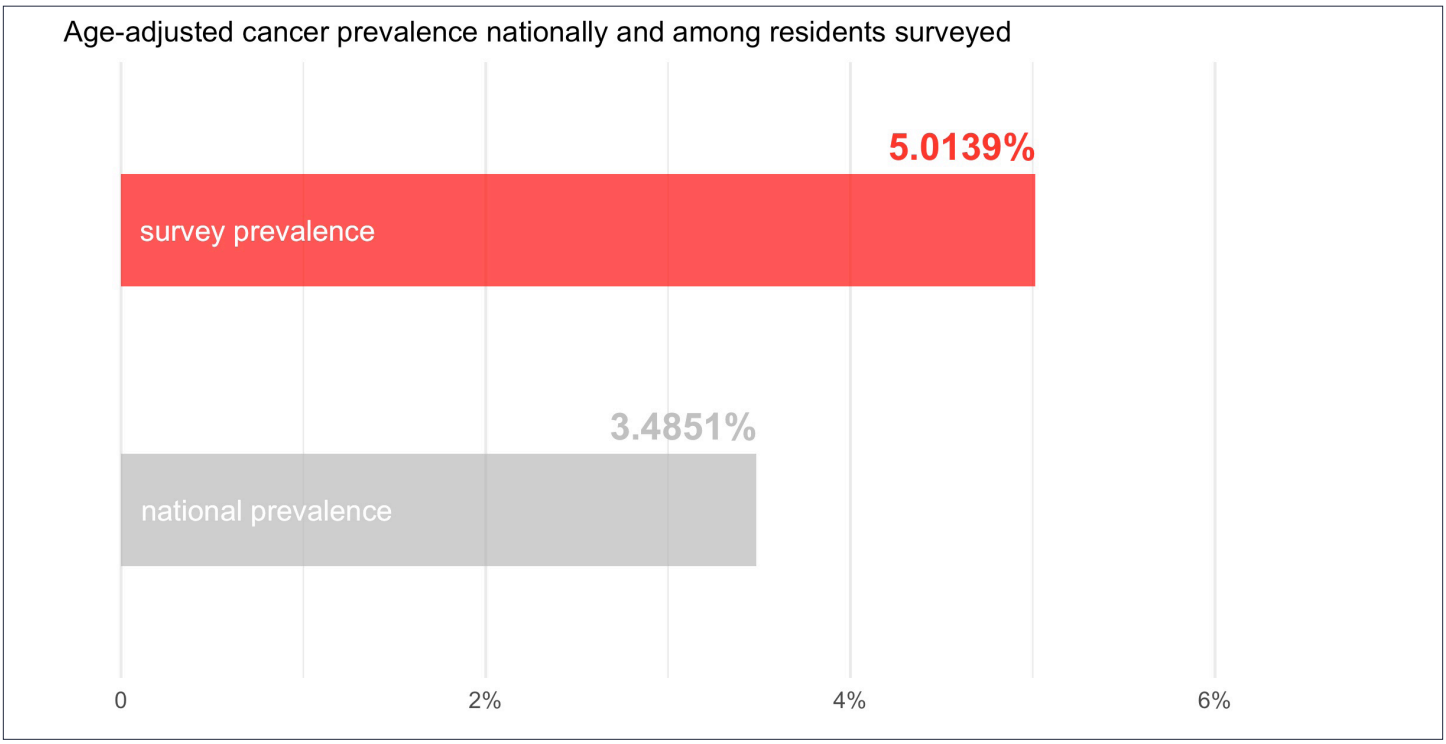
Cancer prevalence among survey respondents is 71% higher than the national rate.



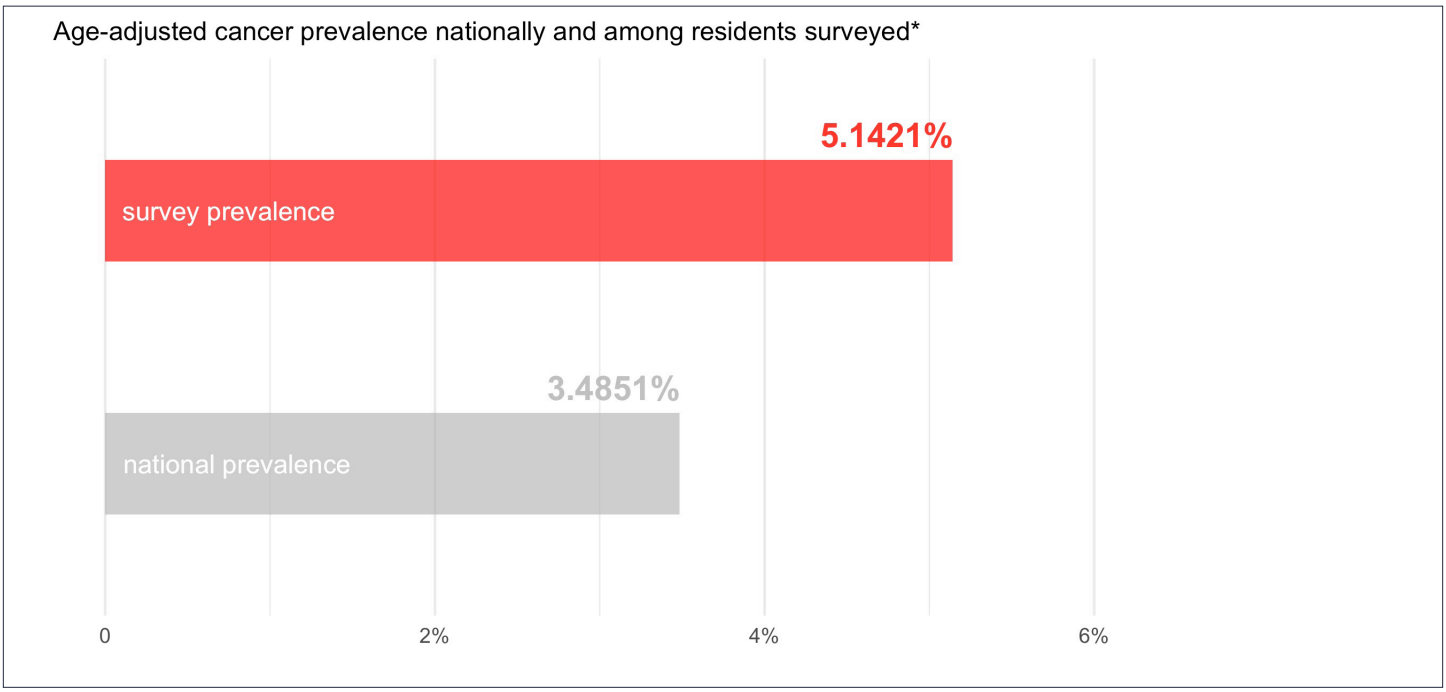
* after removing all respondents who live in households where anyone smokes on a daily basis

Cancer prevalence among survey respondents (with smoking exclusion criterion) is 75% higher than the national rate.

(II) RESIDENT VERSUS NATIONAL CANCER PREVALENCE, without and with smoking exclusion criterion



Cancer prevalence among survey residents is 44% higher than the national rate.



* after removing all residents who live in households where anyone smokes on a daily basis

Cancer prevalence among survey residents (with smoking exclusion criterion) is 48% higher than the national rate.

Introduction to Non-Cancer Health and Pollution Analyses

Introduction to Non-Cancer Health and Pollution Analyses

The EPA's *Toxicological Review of Chloroprene* (2010) contains information on both the carcinogenic and noncarcinogenic effects of chloroprene exposure. Our study results on carcinogenic effects (cancer prevalence) were presented in the previous section. This section presents our results on chloroprene-associated, non-cancer health conditions among residents of the area near the Denka facility.

The EPA's inhalation reference concentration (RfC) for chloroprene is 0.02 milligrams per cubic meter (mg/m3).[1] The RfC, which is typically used in the EPA's non-cancer health assessments, estimates the maximum continuous inhalation exposure to a chemical "without an appreciable risk of deleterious effects during a lifetime."[2] Thus, according to the EPA, humans can be exposed to a maximum chloroprene air concentration of 0.02 mg/m3 without facing an "appreciable risk" of harmful non-cancer health impacts.[3]

We collected data on the following conditions that, according to the EPA, have been linked to chloroprene exposure: headache, dizziness, fatigue, shortness of breath, rapid heart rate, heart palpitations, chest pain, and irritation of the eyes, nose, throat, and skin. These conditions can affect people both short- and long-term following exposure to chloroprene. Below we present our findings on the survey sample prevalence of each of these conditions.

This section also includes our study results on health effects in children. The EPA's 2010 assessment contains evidence suggesting that children are more susceptible than adults to the toxic effects of chloroprene exposure because of their lower capacity to metabolize and excrete chloroprene. Survey respondents were asked whether children in the household suffer from headaches or nosebleeds, since community members frequently cite both of these symptoms as common in children who live and/or attend school in the area near the Denka facility (headaches are also scientifically linked to chloroprene exposure, as noted above).

1 Environmental Protection Agency. *Health Effects Notebook for Hazardous Air Pollutants*, Chloroprene (2-Chloro-1,3 Butadiene). https://www.epa.gov/sites/production/files/2016-10/documents/chloroprene.pdf.

2 Environmental Protection Agency. *Health Effects Notebook Glossary*, RfC (inhalation reference concentration) https://www.epa.gov/haps/health-effects-notebook-glossary.

3 The 0.02 milligram/cubic meter RfC is not to be confused with the EPA's 100-in-1 million and 1-in-1 million cancer-risk based comparison levels for chloroprene, which are 0.2 and 0.002 microgram/cubic meter, respectively. The EPA considers a cancer risk of 100-in-1 million as "the upper limit of acceptability," but aims for control measures that reduce cancer risk to as close to 1-in-1 million as possible.

Finally, this section includes our findings on the frequency and strength of chemical odors in the area, as well as residents' concern about pollution in their community.

Like the cancer prevalence analyses in the previous section, all non-cancer health and pollution analyses were conducted by spatial zone (where "Zone 1" is the area within 1.5 kilometers of the plant and "Zone 2" is the area between 1.5 and 2.5 kilometers from the plant).[4]

We did not use Monte Carlo simulations for our analyses of child health, symptoms, chemical odors, and concern about pollution because we lacked a reliable set of comparable national data broken down by demographic group. Each of these analyses instead includes a bar graph of the data and a corresponding table of values with the total number of people in the pool ("n"), the number of people with the health condition or pollution response (test statistic, or "t"), and the percentage (t/n).

All part-time respondents and residents (defined as those who live in the household for only 1-5 days of the week, inclusive) were eliminated from the data set. The one household that consisted exclusively of part-time residents was also eliminated. After elimination of part-time residents and this household, the total numbers of respondents, residents, and households in the data set were as follows:

	Zone 1	Zone 2	Total
Households	266	271	537
Respondents	263	259	522
Residents	789	754	1,543

We analyzed our survey results on prevalence of rapid pulse/rapid heart rate (tachycardia) using Monte Carlo simulations because tachycardia is a diagnosis for which we had comparable national data broken down by sex (but not race or age). These Monte Carlo analyses of tachycardia prevalence are presented in Part II of the next section.

4 We also analyzed health outcomes in the area within 1.25 kilometers of the plant ("Zone A") and the area between 1.25 and 2.5 kilometers from the plant ("Zone B"). See Appendix for Zone A and Zone B analyses.

Results of Non-Cancer Health and Pollution Analyses

Results of Non-Cancer Health and Pollution Analyses

The follow survey results are presented below:

(I) Child Health:
This section presents our data on headaches and nosebleeds in children.

(II) Rapid Pulse/Heart Rate (Tachycardia) Diagnoses:
This section presents Monte Carlo Analyses of our data on diagnosed[1] tachycardia prevalence among survey respondents and residents.

(III) Symptoms:
This section presents our data on the following chloroprene-associated symptoms experienced by survey respondents (symptoms data were collected for respondents only):

- Chest pain and heart palpitations
- Wheezing and difficulty breathing
- Headaches, dizziness, and lightheadedness
- Eye pain/irritation and watery eyes
- Cough, sneezing, and sore/hoarse throat
- Skin rash/irritation and itchy skin
- Fatigue/lethargy

(IV) Chemical Odors:
This section presents our data on the frequency at which respondents smell chemical odors inside and outside the home. (Chemical odors data were collected for respondents only.)

(V) Concern About Pollution:
This section presents our data on respondents' concern about pollution in their community. (Data on concern about pollution were collected for respondents only.)

1 Survey respondents were asked, "Has a doctor or another health care provider ever told you or anyone else in your household that you or they have rapid pulse or rapid heart rate?"

Our major findings on non-cancer health outcomes among survey respondents are:

- Prevalence of all chloroprene-associated, non-cancer health outcomes in the survey sample was higher in Zone 1 (closer to the plant) than in Zone 2 (farther from the plant).
- Nearly half the children in the households surveyed in Zone 1 suffer from headaches, nosebleeds, or both.
- Prevalence of rapid pulse/rapid heart rate (tachycardia) diagnoses in the survey sample is unusually high—multiple times higher than expected and virtually impossible to be due to chance.
- Nearly 40% of Zone 1 respondents regularly experience chest pain, heart palpitations, or both.
- One-third of Zone 1 respondents regularly experience wheezing and/or difficulty breathing.
- More than half of Zone 1 respondents regularly experience headaches, dizziness, and/or lightheadedness.
- Nearly half of Zone 1 respondents regularly experience eye pain/irritation and/or watery eyes.
- More than 40% of Zone 1 respondents experience cough, sneezing, and/or sore/hoarse throat most of the time.
- More than one-third of Zone 1 respondents regularly experience skin rash/irritation and/or itchy skin.
- Nearly 30% of Zone 1 respondents experience fatigue/lethargy most of the time.

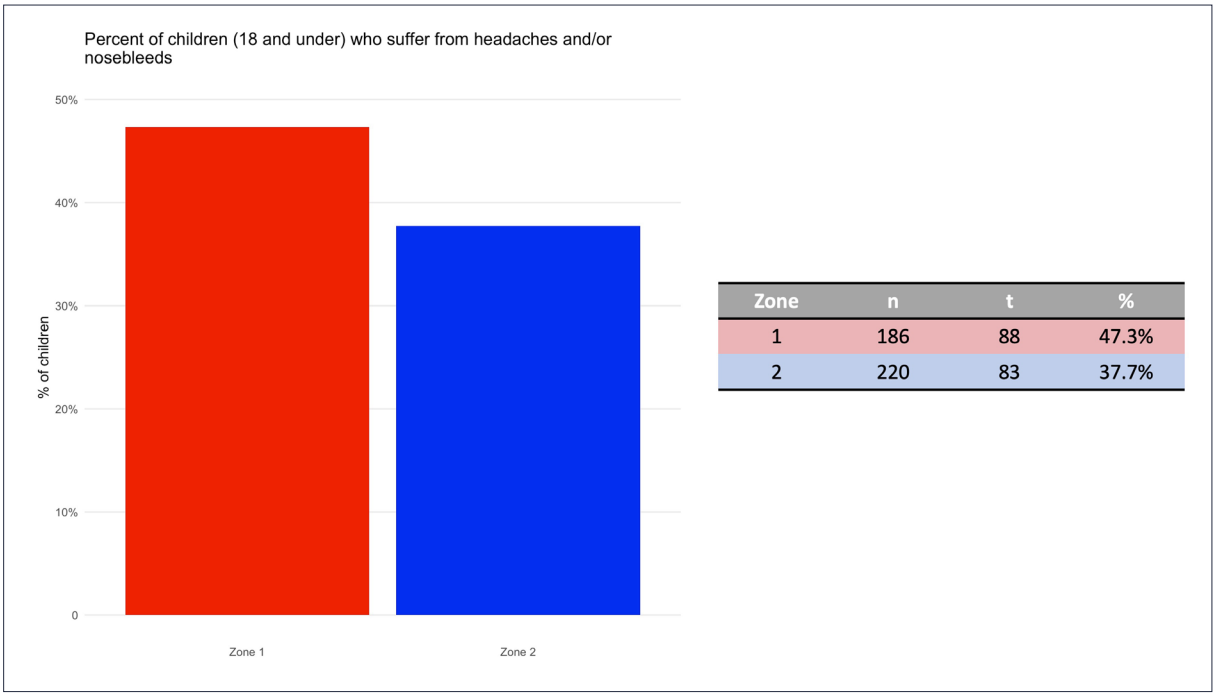
We also found that nearly half of Zone 1 respondents smell chemical odors *inside* their homes at least a few times per month; over half of Zone 1 respondents smell chemical odors outside their homes at least a few times per *week*; and over three-fourths of Zone 1 respondents smell chemical odors outside their homes at least a few times per month.

84% of Zone 1 respondents reported that they are "extremely concerned" about pollution in their community.

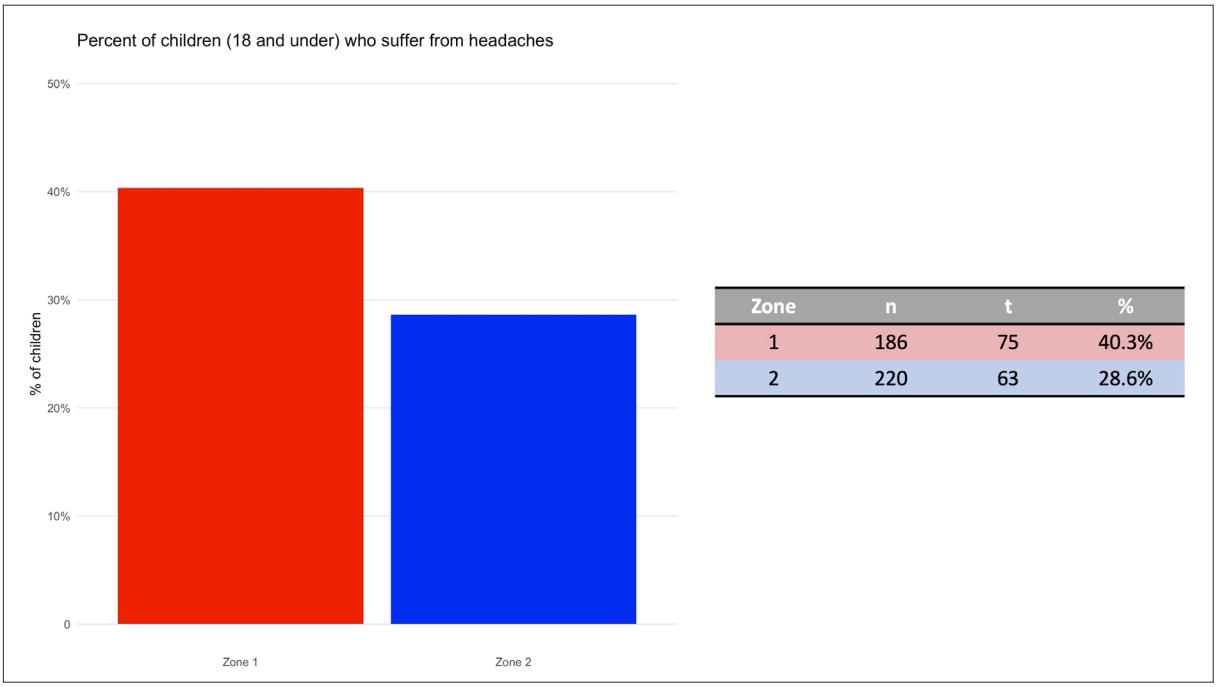
(I) CHILD HEALTH

- **Nearly half the children in the households surveyed in Zone 1 suffer from headaches, nosebleeds, or both.**
- **Over 40% of the childen in the households surveyed in Zone 1 suffer from headaches, which are associated with chloroprene exposure.**

FIGURE 6.1 (RIGHT)



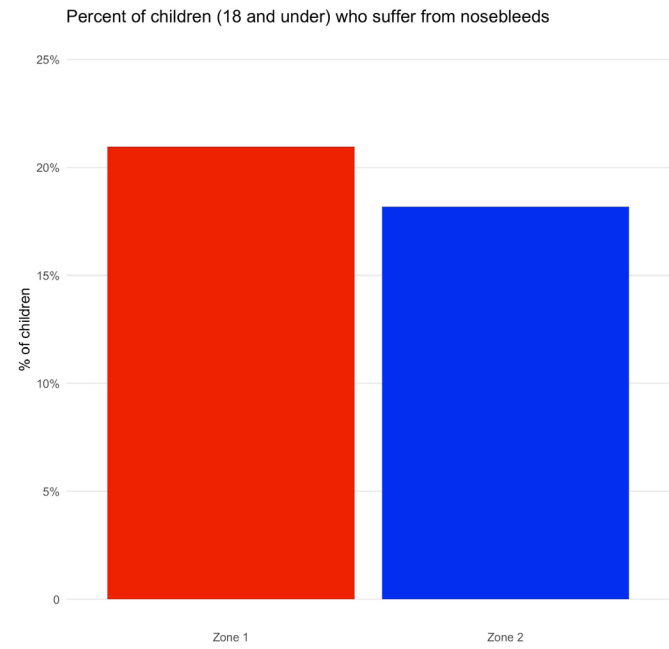
Zone	n	t	%
1	186	88	47.3%
2	220	83	37.7%



Zone	n	t	%
1	186	75	40.3%
2	220	63	28.6%

FIGURE 6.2 (LEFT)

FIGURE 6.3 (RIGHT)



Zone	n	t	%
1	186	39	21%
2	220	40	18.2%

(II) RAPID PULSE/RAPID HEART RATE (TACHYCHARDIA) DIAGNOSES

- **Prevalence of rapid pulse/rapid heart rate (tachycardia) diagnoses in the survey sample is unusually high—multiple times higher than expected and virtually impossible to be due to chance.**

The next series of visualizations (**Figures 7.1—7.8**) are Monte Carlo analyses of diagnosed rapid pulse/rapid heart rate (tachycardia) among survey respondents and residents. Survey respondents were asked, "Has a doctor or another health care provider ever told you or anyone else in your household that you or they have rapid pulse or rapid heart rate?"

According to the CDC's National Center for Health Statistics, there are two definitions of tachycardia. By the clinical consensus definition (resting pulse rate greater than 100 beats/minute), prevalence of tachycardia is 1.3% in adult males and 1.9% in adult females. By the revised clinical guideline (resting pulse rate greater than 90 beats/minute), prevalence of tachycardia is 5.2% in adult males and 8.4% in adult females. The National Center for Health Statistics defines "adult" as 20 years or older.[1]

Since national prevalence rates for tachycardia were available by sex only, we could consider only sex in the Monte Carlo analyses below. We eliminated from the data set one respondent/resident for whom we did not have sex information.

1 Ostchega Y, Porter KS, Hughes J, Dillion CF, Nwankwo T. Resting pulse reference data for children, adolescents, and adults: United States, 1999—2008. National health statistics reports; no 41. Hyattsville, MD: National Center for Health Statistics. 2011.

Since national prevalence rates by race were not available, we could not consider race in the tachycardia analyses below. Respondents and residents for whom we did not have race information were *not* eliminated from the data set as a result.

Since national prevalence rates by age were not available, we could not consider age in the tachycardia analyses below, except by limiting our analyses to adults (20 years or older) only. To limit our analyses to adults, we eliminated from the data set (1) all respondents and residents for whom we did not have age information, and (2) all respondents and residents who were 19 years of age or younger.

After eliminations, the total numbers of respondents and residents included in our Monte Carlo analyses of tachycardia prevalence were as follows:

	Zone 1	Zone 2	Total
Respondents	259	252	511
Residents	594	527	1,121

The graphs below show expected tachycardia prevalence among members of a simulated population with the same sex demographics as our adult survey sample. The graphs also show actual tachycardia prevalence among adult respondents and residents, and the likelihood of this prevalence.

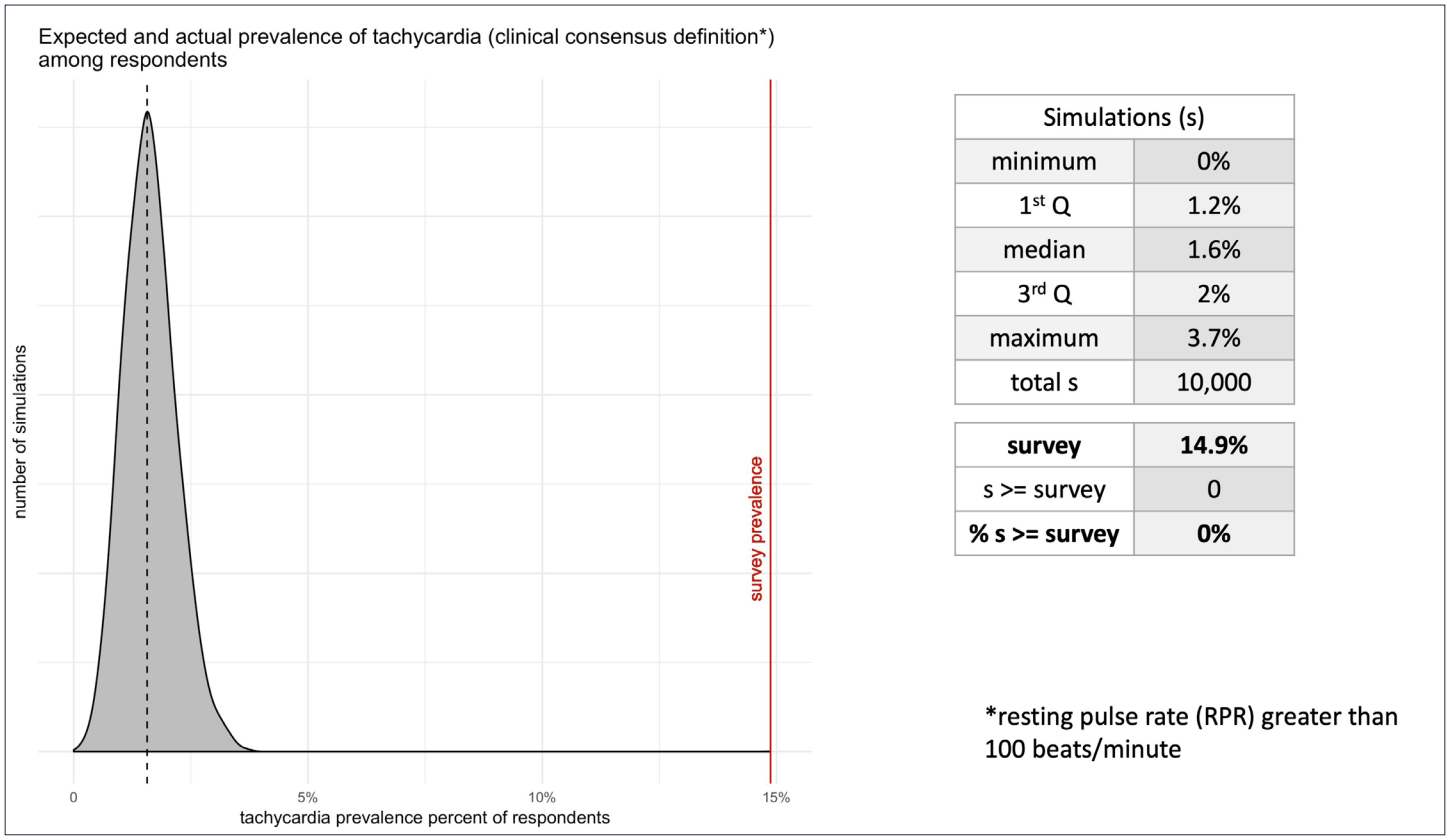
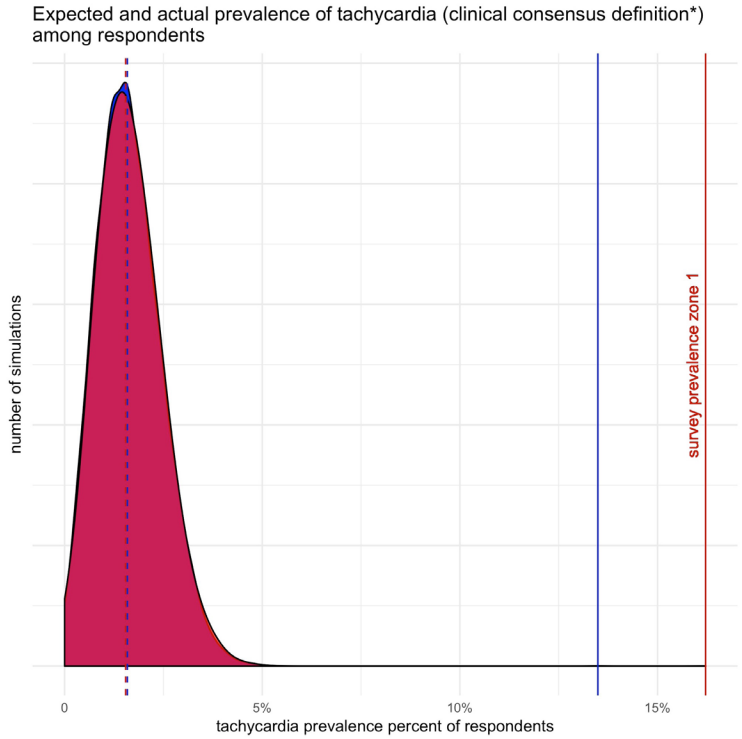


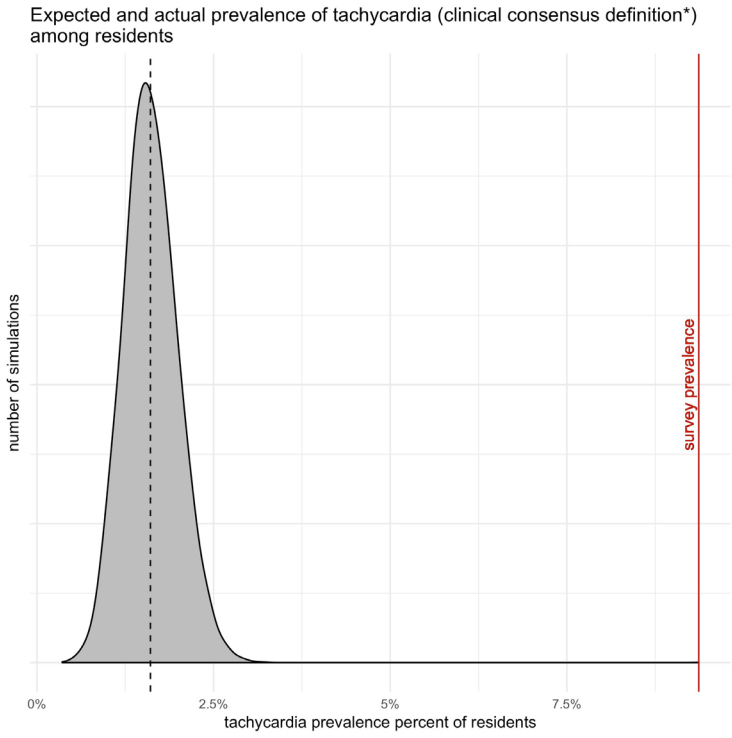
FIGURE 7.1 (ABOVE): Extremely statistically significant p-value of 0%, indicating virtual impossibility that 14.9% prevalence of tachycardia (by clinical consensus definition) among respondents is due to chance



Simulations (s)		
Zone	1	2
minimum	0%	0%
1st Q	1.2%	1.2%
median	1.5%	1.6%
3rd Q	2.3%	2%
maximum	5.4%	5.2%
total s	10,000	10,000
survey	**16.2%**	**13.5%**
s >= survey	0	0
% s >= survey	**0%**	**0%**

*resting pulse rate (RPR) greater than 100 beats/minute

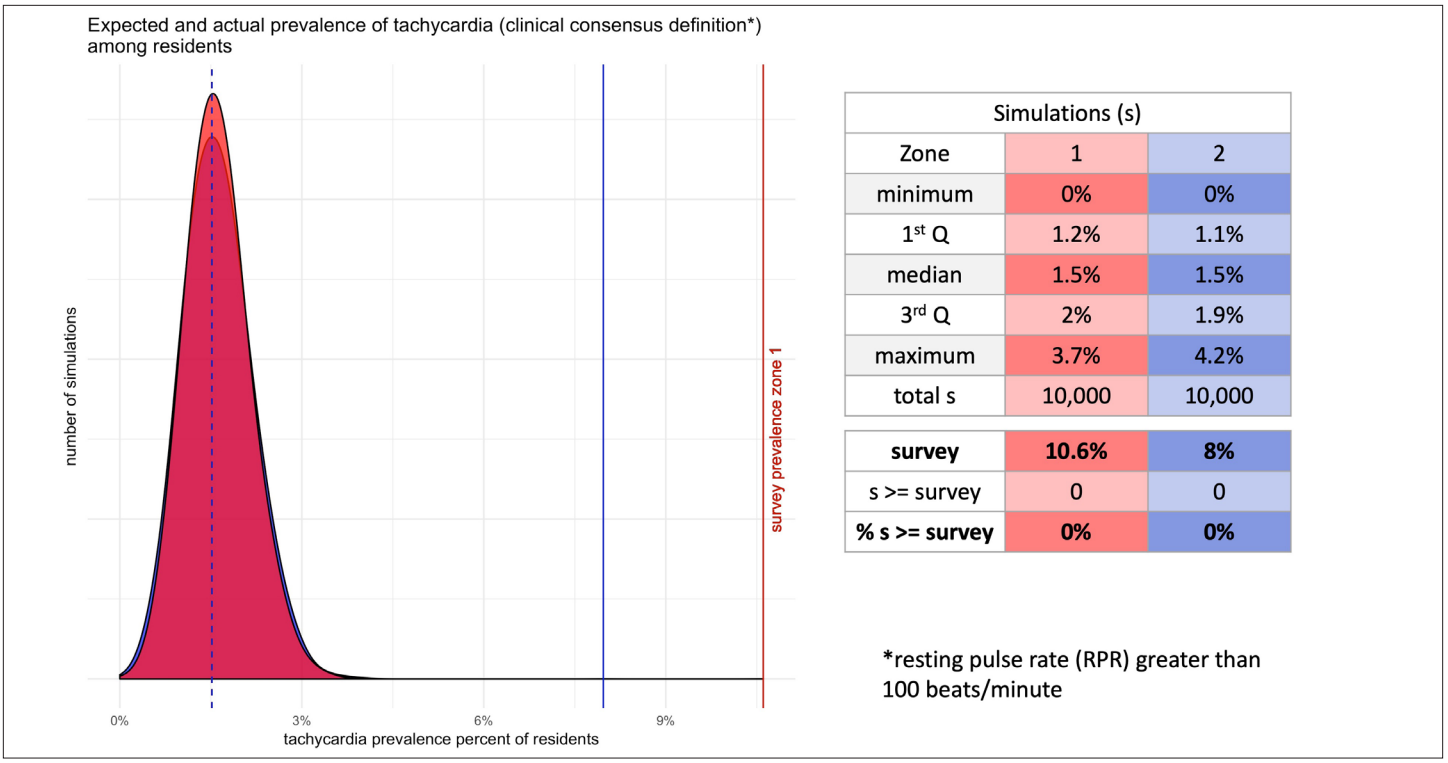
FIGURE 7.2 (ABOVE): Extremely statistically significant p-values of 0% for both Zone 1 and Zone 2, indicating virtual impossibility that 16.2% and 13.5% prevalence of tachycardia (by clinical consensus definition) among Zone 1 and Zone 2 respondents, respectively, is due to chance



Simulations (s)	
minimum	0.4%
1st Q	1.3%
median	1.6%
3rd Q	1.9%
maximum	3.2%
total s	10,000
survey	**9.4%**
s >= survey	0
% s >= survey	**0%**

*resting pulse rate (RPR) greater than 100 beats/minute

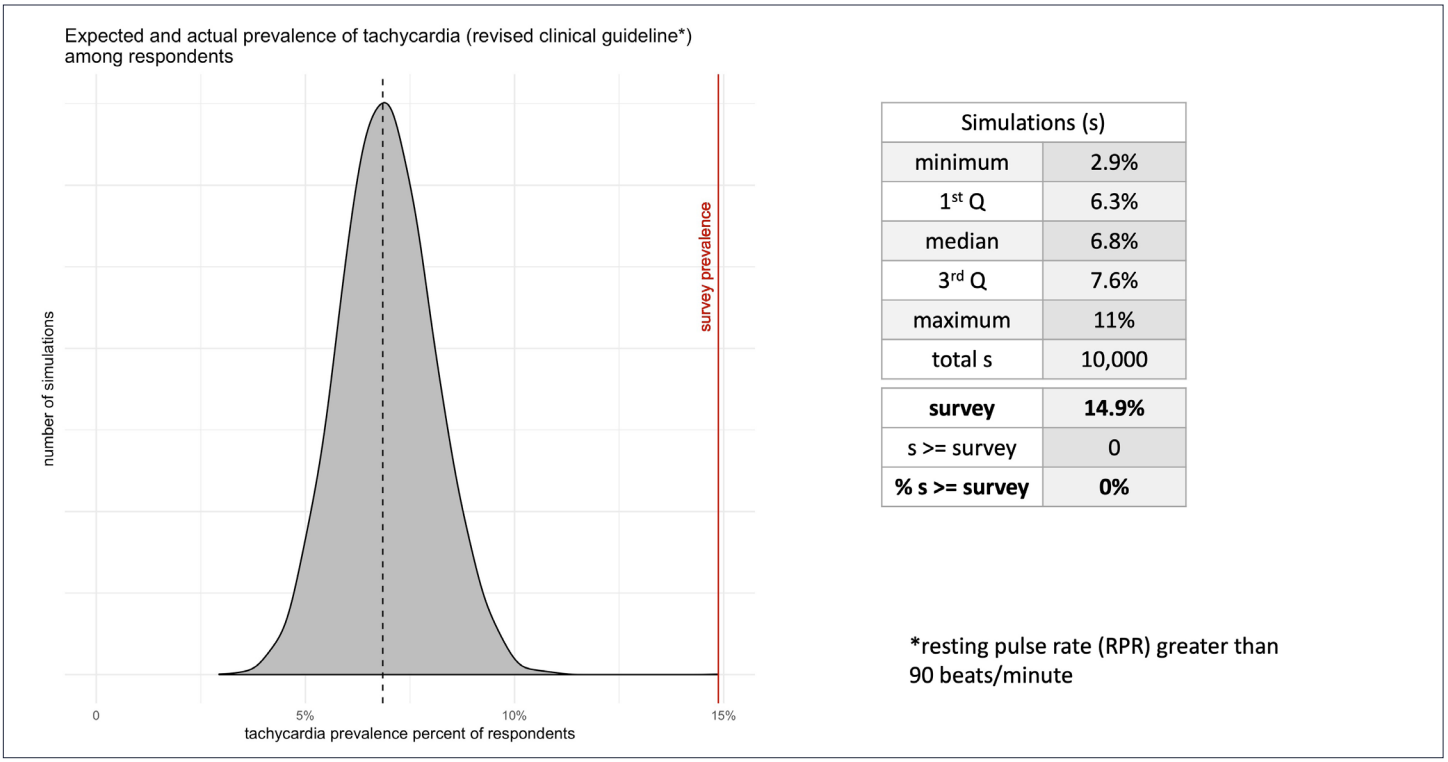
FIGURE 7.3 (ABOVE): Extremely statistically significant p-value of 0%, indicating virtual impossibility that 9.4% prevalence of tachycardia (by clinical consensus definition) among residents is due to chance



Simulations (s)		
Zone	1	2
minimum	0%	0%
1st Q	1.2%	1.1%
median	1.5%	1.5%
3rd Q	2%	1.9%
maximum	3.7%	4.2%
total s	10,000	10,000
survey	**10.6%**	**8%**
s >= survey	0	0
% s >= survey	**0%**	**0%**

*resting pulse rate (RPR) greater than 100 beats/minute

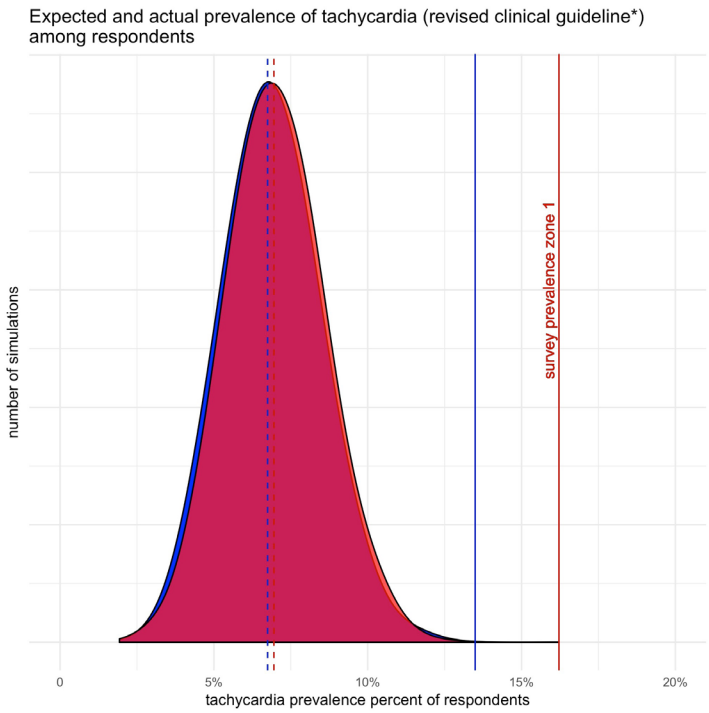
FIGURE 7.4 (ABOVE): Extremely statistically significant p-values of 0% for both Zone 1 and Zone 2, indicating virtual impossibility that 10.6% and 8% prevalence of tachycardia (by clinical consensus definition) among Zone 1 and Zone 2 residents, respectively, is due to chance



Simulations (s)	
minimum	2.9%
1st Q	6.3%
median	6.8%
3rd Q	7.6%
maximum	11%
total s	10,000
survey	**14.9%**
s >= survey	0
% s >= survey	**0%**

*resting pulse rate (RPR) greater than 90 beats/minute

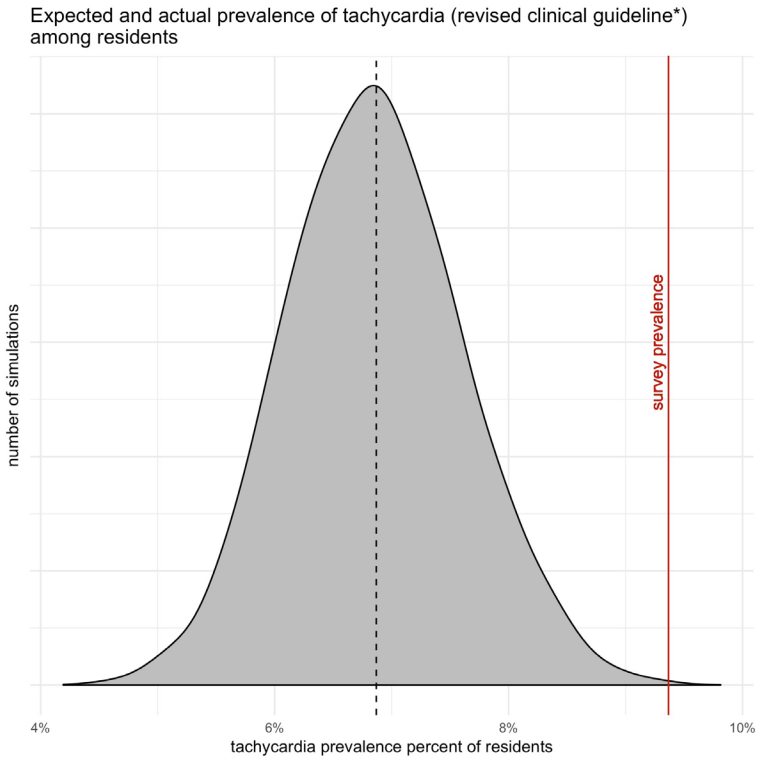
FIGURE 7.5 (ABOVE): Extremely statistically significant p-value of 0%, indicating virtual impossibility that 14.9% prevalence of tachycardia (by revised clinical guideline) among respondents is due to chance



Simulations (s)		
Zone	1	2
minimum	1.9%	2%
1st Q	5.8%	6%
median	6.9%	6.7%
3rd Q	8.1%	7.9%
maximum	13.1%	13.5%
total s	10,000	10,000
survey	**16.2%**	**13.5%**
s >= survey	0	1
% s >= survey	**0%**	**0.01%**

*resting pulse rate (RPR) greater than 90 beats/minute

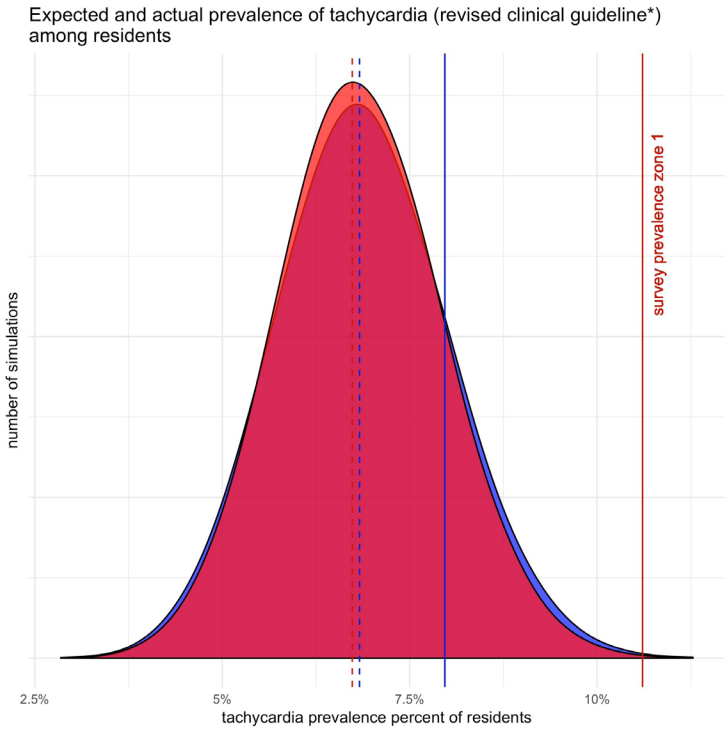
FIGURE 7.6 (ABOVE): **Extremely statistically significant p-values of 0% for Zone 1 and 0.01% for Zone 2, indicating virtual impossibility that 16.2% and 13.5% prevalence of tachycardia (by revised clinical guideline) among Zone 1 and Zone 2 respondents, respectively, is due to chance**



Simulations (s)	
minimum	4.2%
1st Q	6.3%
median	6.9%
3rd Q	7.3%
maximum	9.8%
total s	10,000
survey	**9.4%**
s >= survey	6
% s >= survey	**0.06%**

*resting pulse rate (RPR) greater than 90 beats/minute

FIGURE 7.7 (ABOVE): **Extremely statistically significant p-value of 0.06%, indicating strong unlikelihood that 9.4% prevalence of tachycardia (by revised clinical guideline) among residents is due to chance**



Simulations (s)		
Zone	1	2
minimum	3.4%	2.8%
1st Q	6.1%	6.1%
median	6.7%	6.8%
3rd Q	7.6%	7.6%
maximum	11.3%	11.2%
total s	10,000	10,000

survey	**10.6%**	**8%**
s >= survey	3	1812
% s >= survey	**0.03%**	**18.12%**

*resting pulse rate (RPR) greater than 90 beats/minute

FIGURE 7.8 (ABOVE): Extremely statistically significant p-value of 0.03% for Zone 1, indicating strong unlikelihood that 10.6% prevalence of tachycardia (by revised clinical guideline) among Zone 1 residents is due to chance

---

Figures 7.1—7.4 show tachycardia prevalence by the clinical consensus definition (resting pulse rate greater than 100 beats/minute), which is the generally accepted definition of tachycardia.

In **Figures 7.1 and 7.3**, the medians of both curves (depicted by the dotted vertical lines) are 1.6%, which means that there is a 50% chance that tachycardia prevalence in a simulated population with the same sex makeup as our respondent and resident groups would be greater than 1.6%. As the corresponding tables show, the third quartile tachycardia prevalence is around 2% for both respondents and residents; that is, there is only a 25% chance that tachycardia prevalence (by the clinical consensus definition) would be greater than 2%.

Actual tachycardia prevalence in our respondent and resident groups (depicted by the solid vertical lines) is 14.9% and 9.4%, respectively. Both of these values are several times greater than the *maximum* simulated values. The probability that tachycardia prevalence would be this high in populations with the same sex demographics as our respondent and resident groups (the "p-value")

is 0%. In other words, **it is virtually impossible that tachycardia prevalence this high among survey respondents and residents could be due to chance.**

In **Figures 7.2 and 7.4**, there are two distributions—one for Zone 1 (red) and another for Zone 2 (blue). Tachycardia prevalence among respondents and residents in Zone 1 is 16.2% and 10.6%, respectively. Both of these values are several times greater than the *maximum* simulated values for Zone 1. The probability that tachycardia prevalence would be this high in populations with the same sex demographics as our Zone 1 respondent and resident groups (the "p-value") is 0%. In other words, **it is virtually impossible that tachycardia prevalence this high among Zone 1 survey respondents and residents could be due to chance.**

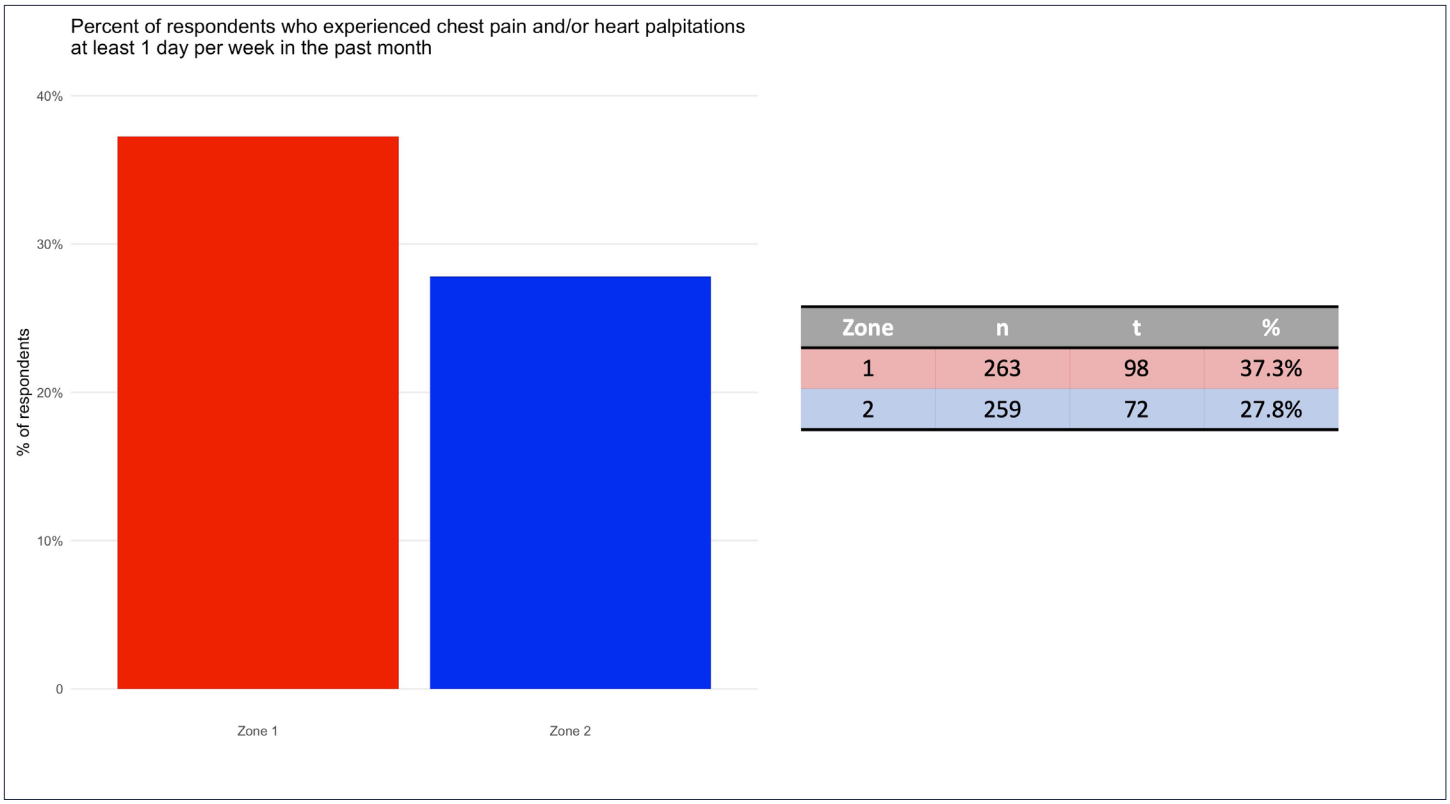
In Zone 2, tachycardia prevalence drops to 13.5% and 8% among respondents and residents, respectively. **This drop in tachycardia prevalence from Zone 1 to Zone 2 indicates that there is an association between closer proximity to the Denka facility and higher tachycardia prevalence among respondents and residents surveyed.** That said, both of these Zone 2 values are still several times greater than the maximum simulated values for Zone 2. Furthermore, the probability that tachycardia prevalence would be this high in populations with the same sex demographics as our Zone 2 respondent and resident groups (the "p-value") is 0% (just as it was for Zone 1 respondents and residents). In other words, **it is virtually impossible that tachycardia prevalence this high among Zone 2 survey respondents and residents could be due to chance.**

Figures 7.5—7.8 show tachycardia prevalence by the revised clinical guideline (resting pulse rate greater than 90 beats/minute). Because the revised clinical guideline threshold for tachycardia is lower, national rates of tachycardia under the revised clinical guideline are significantly higher than they are under the clinical consensus definition. Even compared to these higher national rates, tachycardia prevalence among survey respondents and residents under the revised clinical guideline is unusually high. The p-value for overall respondent prevalence (**Figure 7.5**) is 0% and the p-value for overall resident prevalence (**Figure 7.7**) is 0.06%. The p-value for Zone 1 respondent prevalence (**Figure 7.6**) is 0% and the p-value for Zone 1 resident prevalence (**Figure 7.8**) is 0.03%. The p-value for Zone 2 respondent prevalence is 0.01% (**Figure 7.6**) and the p-value for Zone 2 resident prevalence (**Figure 7.8**) is 18.12%. In other words, **among all groups except Zone 2 residents, it is either virtually impossible or extremely unlikely that high rates of tachycardia (by the revised clinical guideline) could be due to chance.**

(III) SYMPTOMS

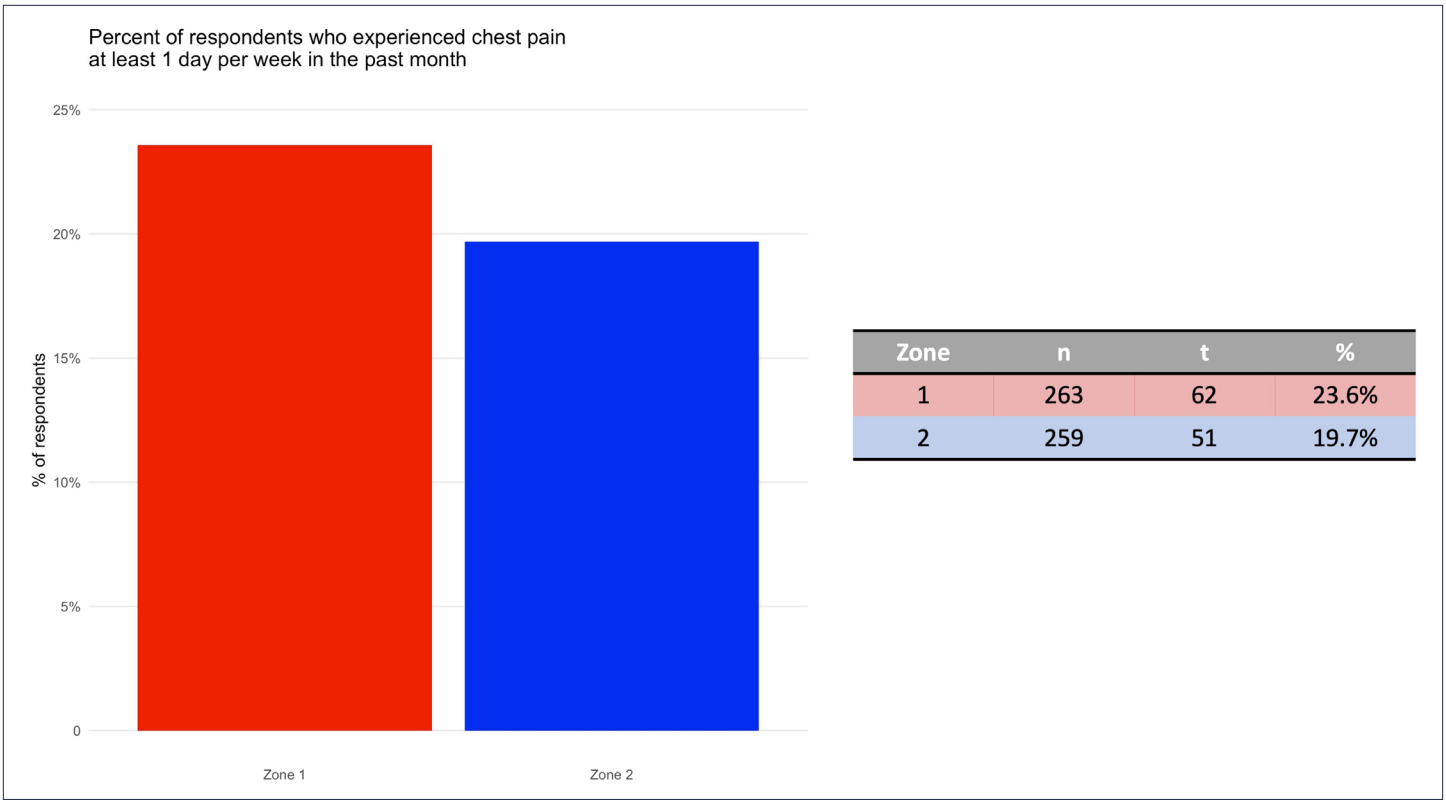
Chest pain and heart palpitations

- **Nearly 40% of Zone 1 respondents regularly experience chest pain, heart palpitations, or both.**



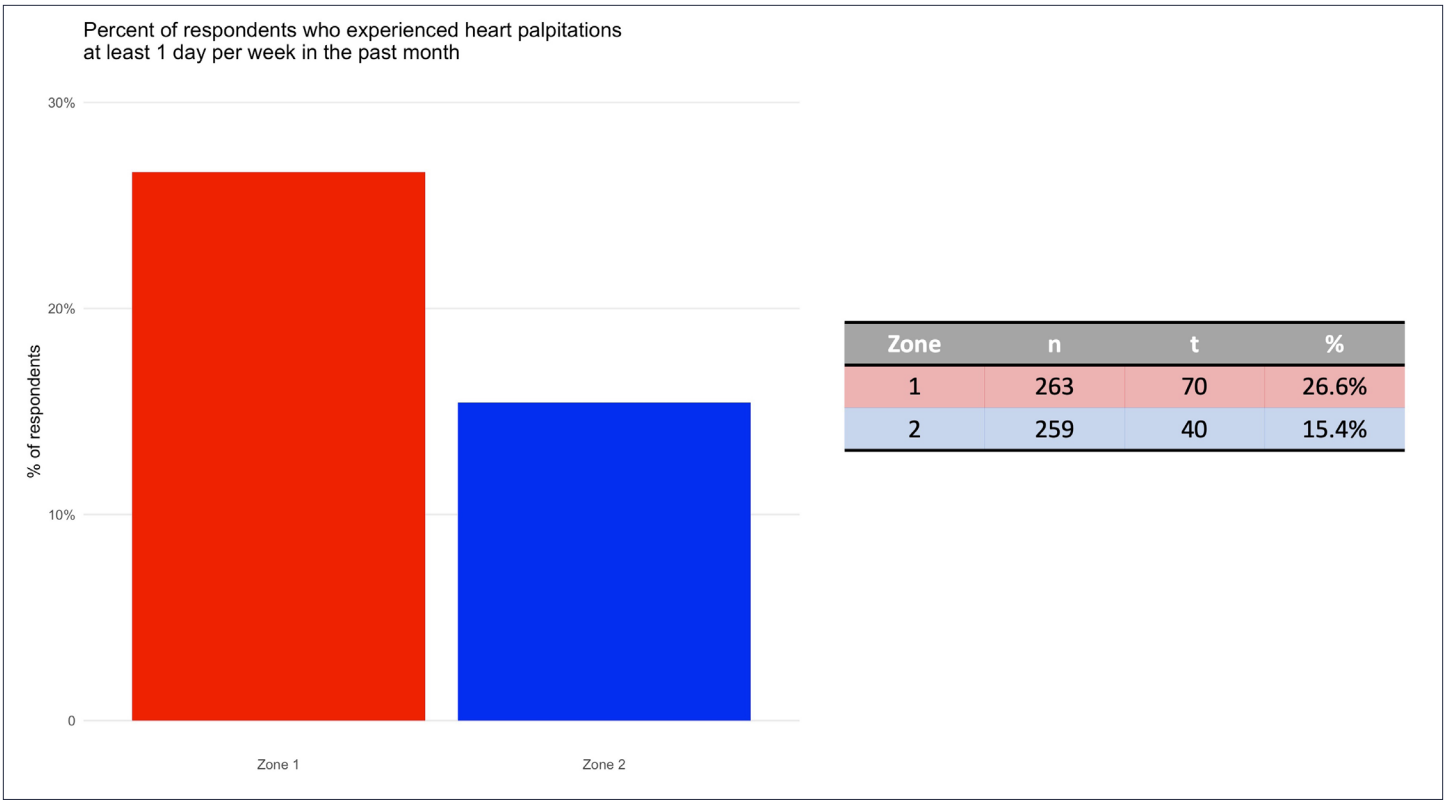
Zone	n	t	%
1	263	98	37.3%
2	259	72	27.8%

FIGURE 8.1 (ABOVE)



Zone	n	t	%
1	263	62	23.6%
2	259	51	19.7%

FIGURE 8.2 (ABOVE)

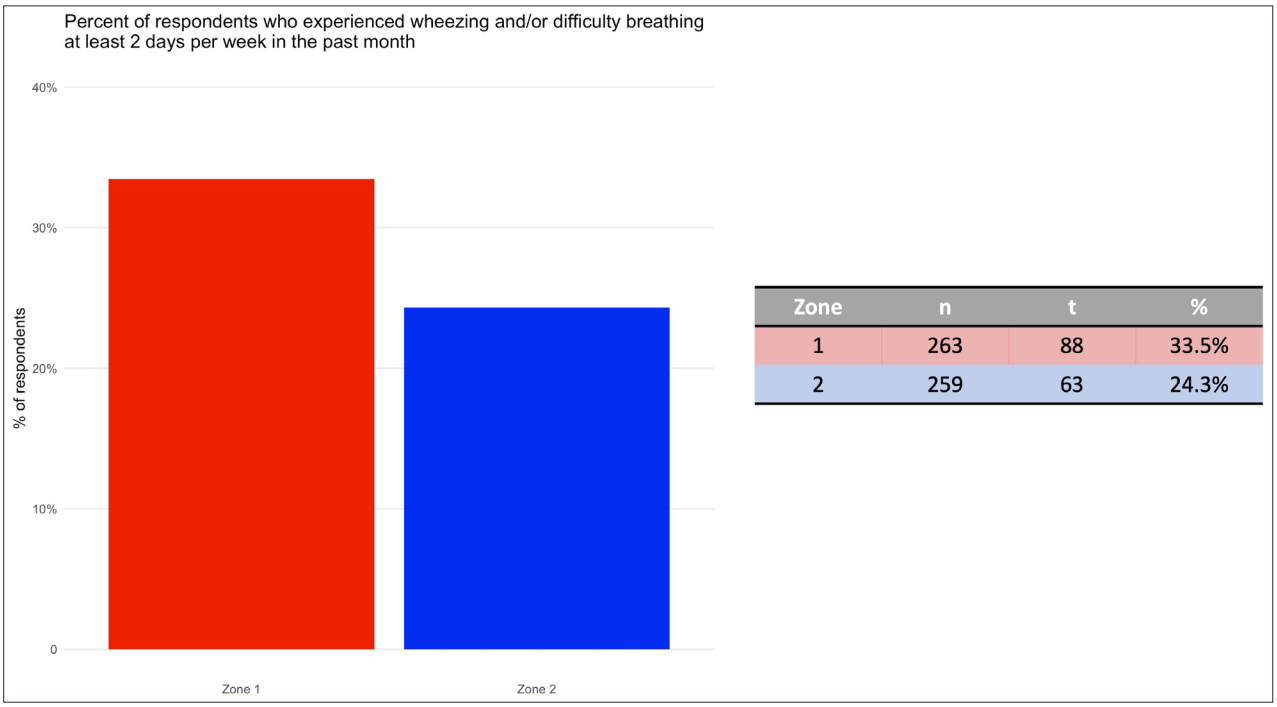


Zone	n	t	%
1	263	70	26.6%
2	259	40	15.4%

FIGURE 8.3 (ABOVE)

Wheezing and difficulty breathing

- **One-third of Zone 1 respondents regularly experience wheezing and/or difficulty breathing.**

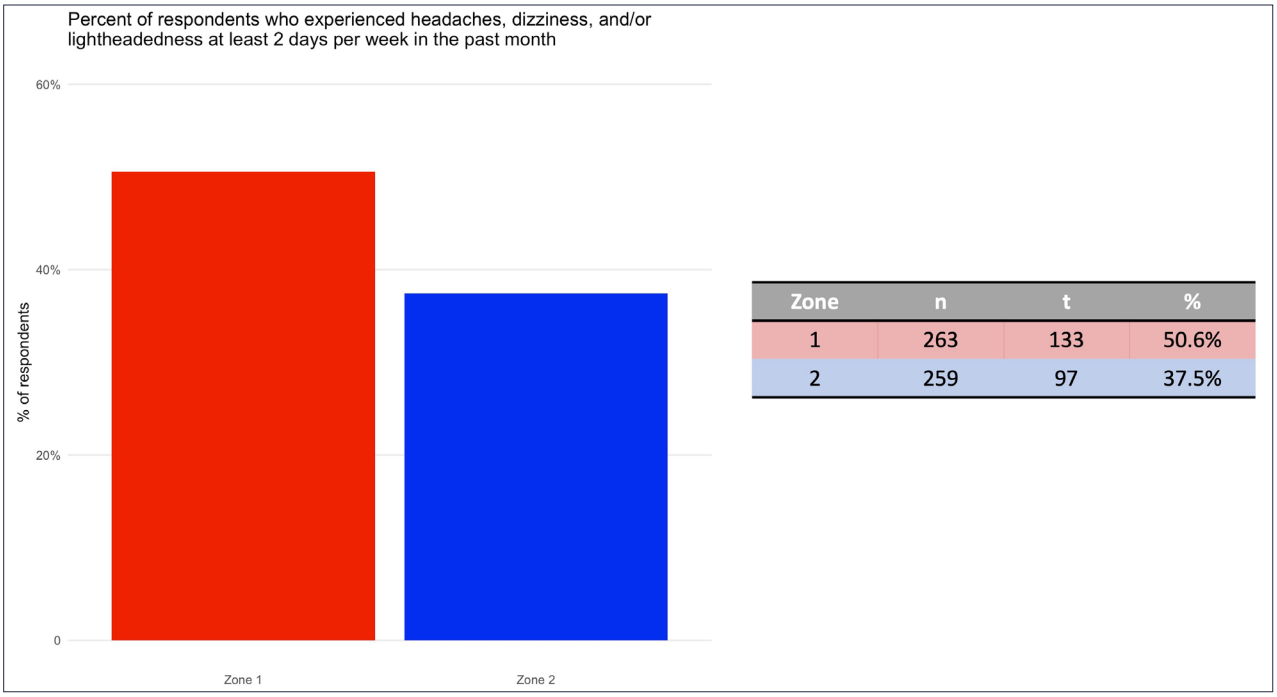


Zone	n	t	%
1	263	88	33.5%
2	259	63	24.3%

FIGURE 8.4 (LEFT)

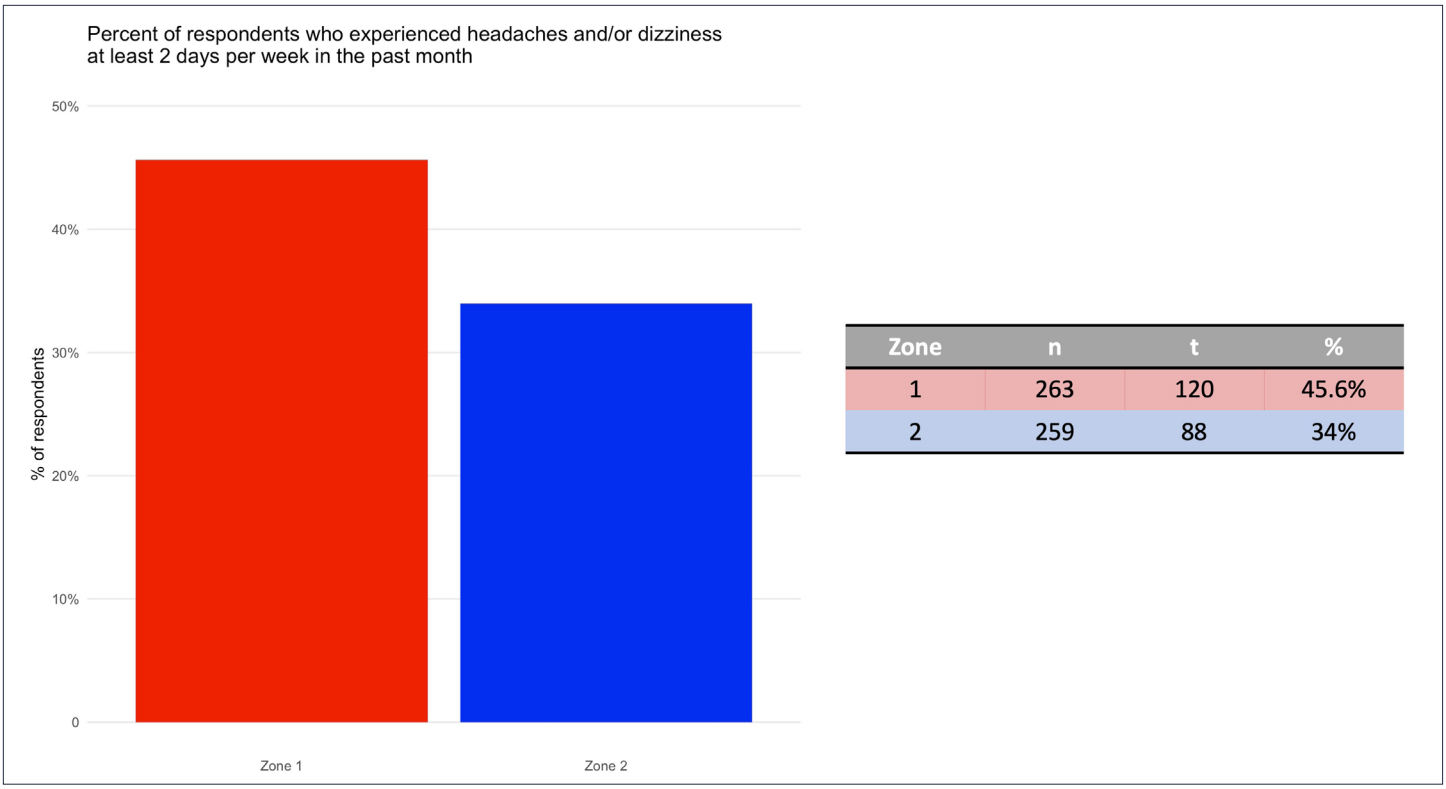
Headaches, dizziness, and lightheadedness

- **More than half of Zone 1 respondents regularly experience headaches, dizziness, and/or lightheadedness.**



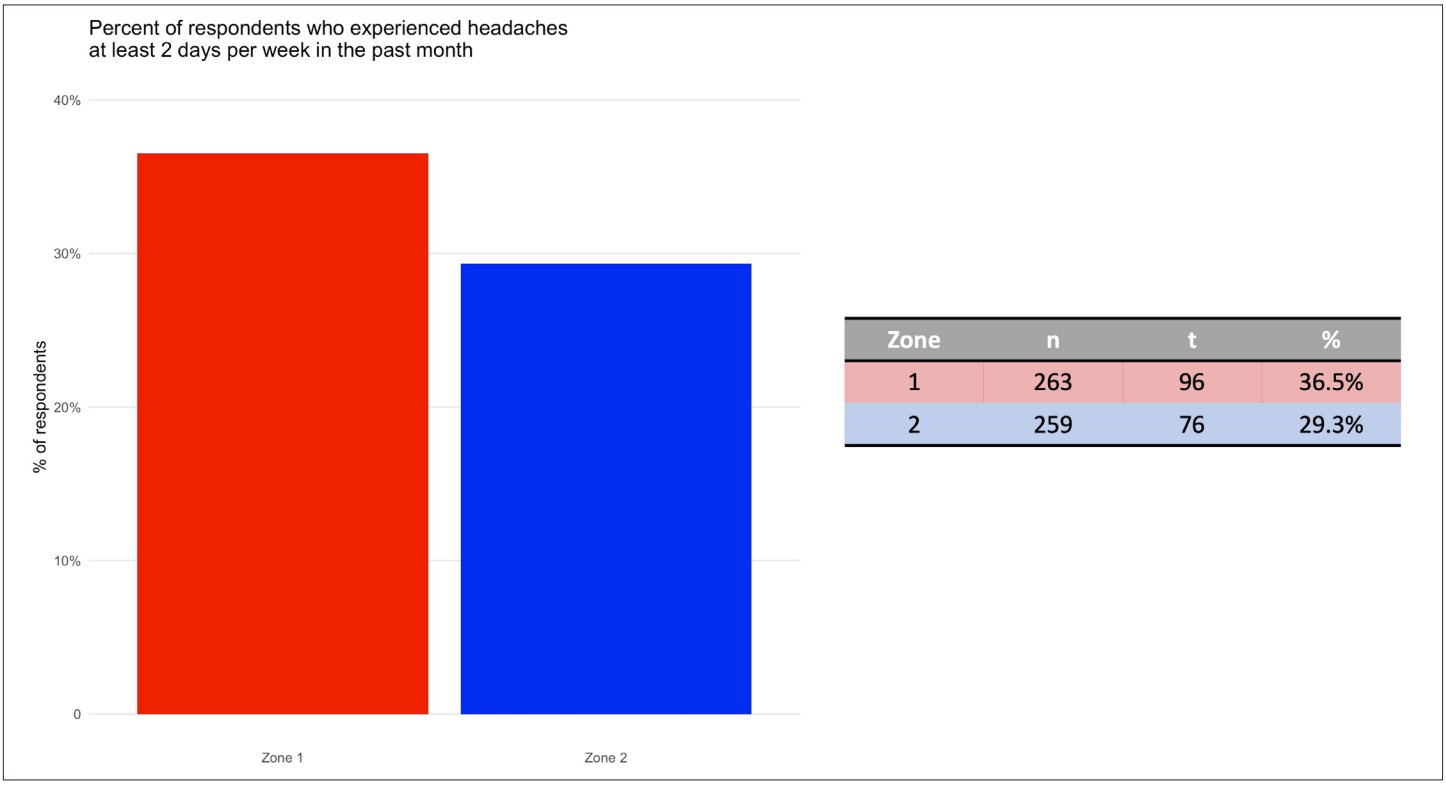
Zone	n	t	%
1	263	133	50.6%
2	259	97	37.5%

FIGURE 8.5 (LEFT)



Zone	n	t	%
1	263	120	45.6%
2	259	88	34%

FIGURE 8.6 (ABOVE)

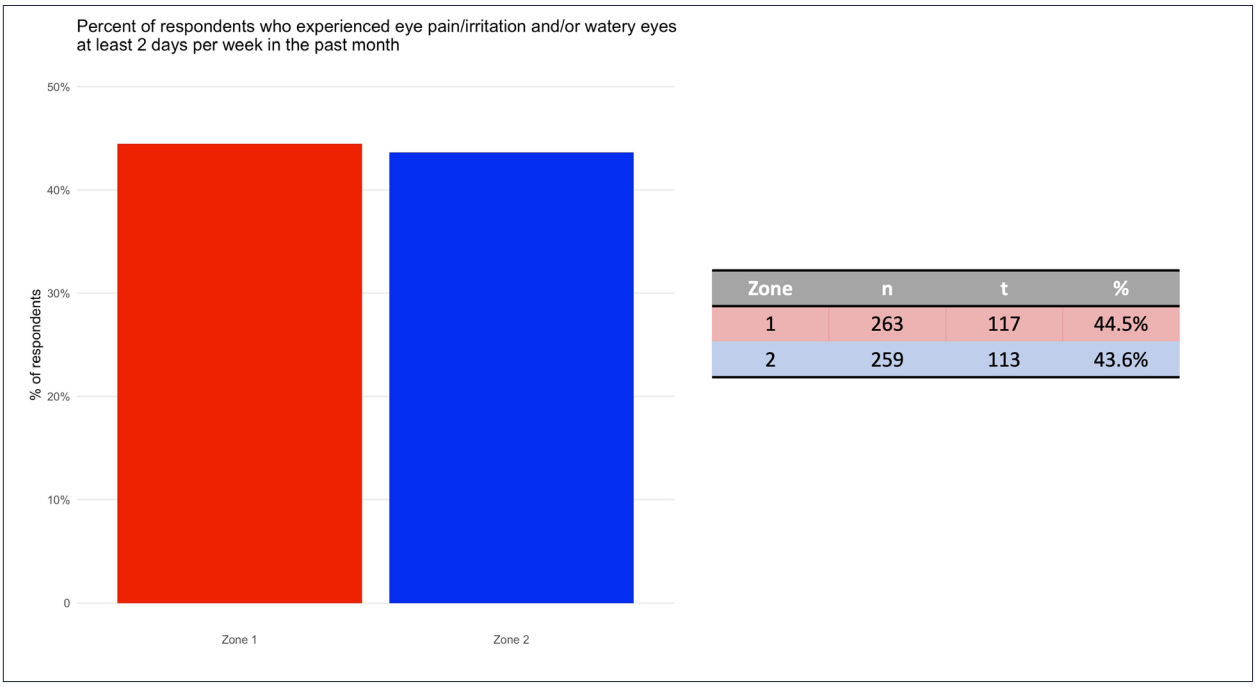


Zone	n	t	%
1	263	96	36.5%
2	259	76	29.3%

FIGURE 8.7 (ABOVE)

Eye pain/irritation and watery eyes

- **Nearly half of Zone 1 respondents regularly experience eye pain/irritation and/or watery eyes.**

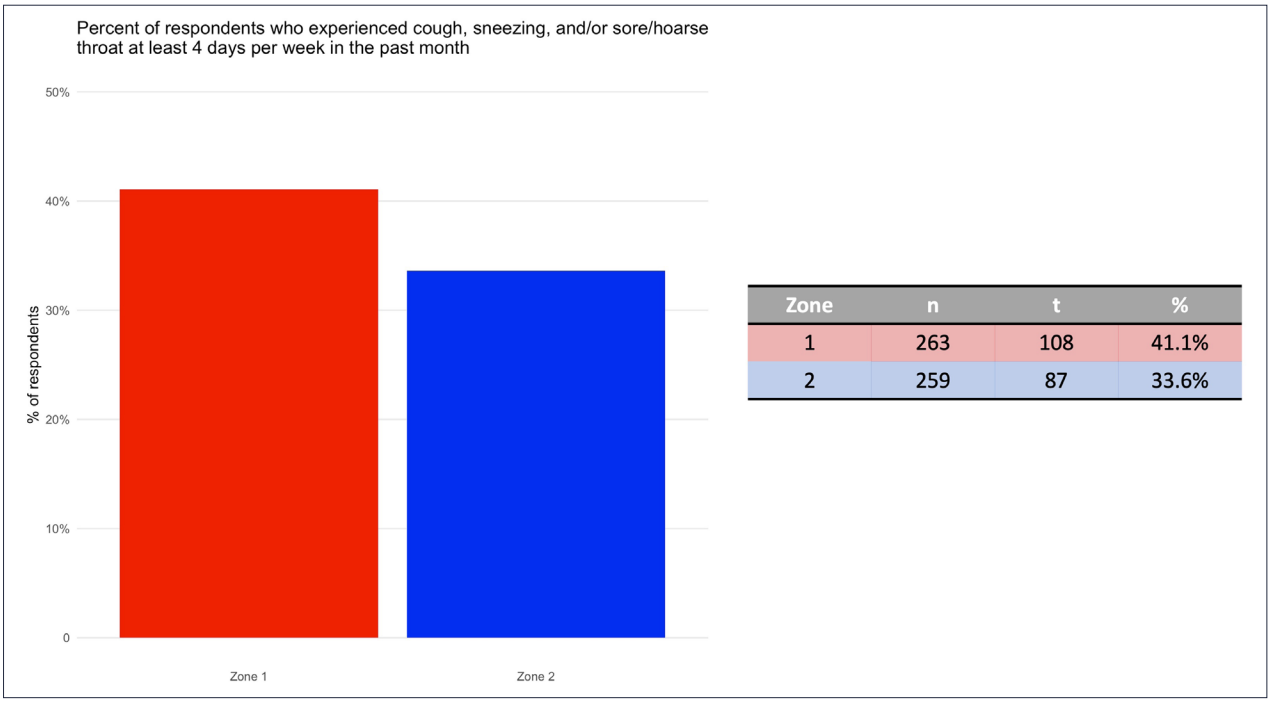


Zone	n	t	%
1	263	117	44.5%
2	259	113	43.6%

FIGURE 8.8 (LEFT)

Cough, sneezing, and sore/hoarse throat

- **More than 40% of Zone 1 respondents experience cough, sneezing, and/or sore/hoarse throat most of the time.**

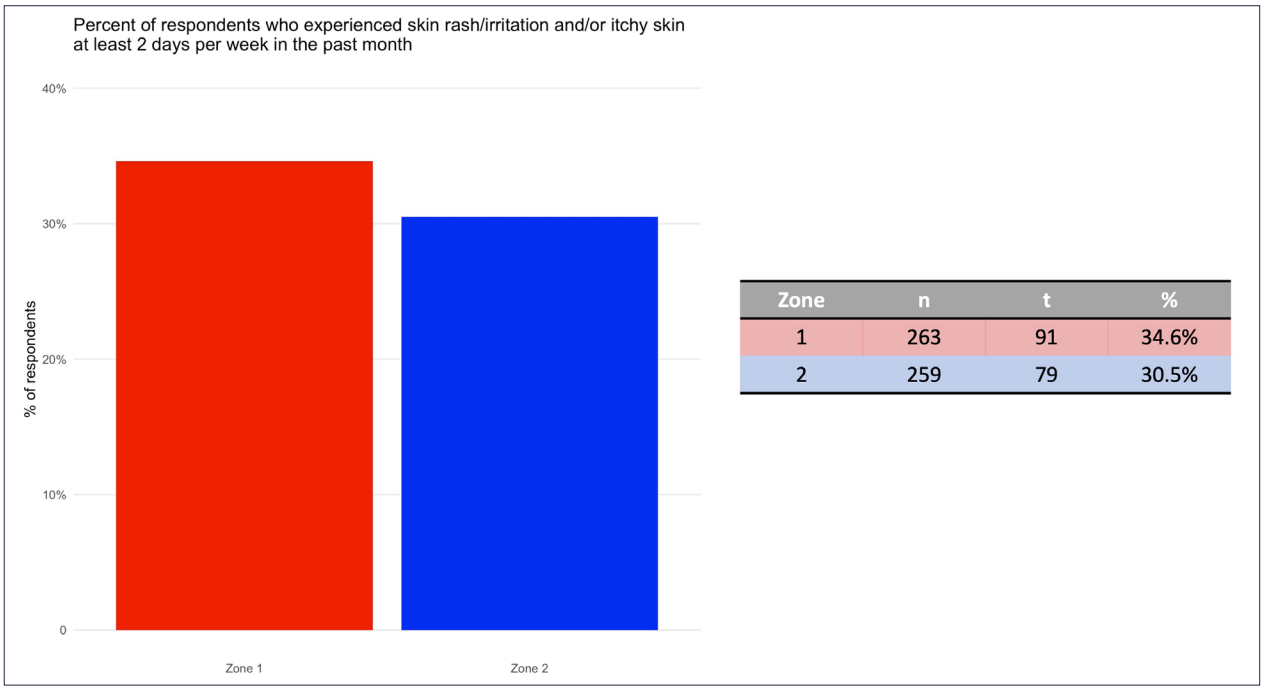


Zone	n	t	%
1	263	108	41.1%
2	259	87	33.6%

FIGURE 8.9 (LEFT)

Skin rash/irritation and itchy skin

- **More than one-third of Zone 1 respondents regularly experience skin rash/irritation and/or itchy skin.**

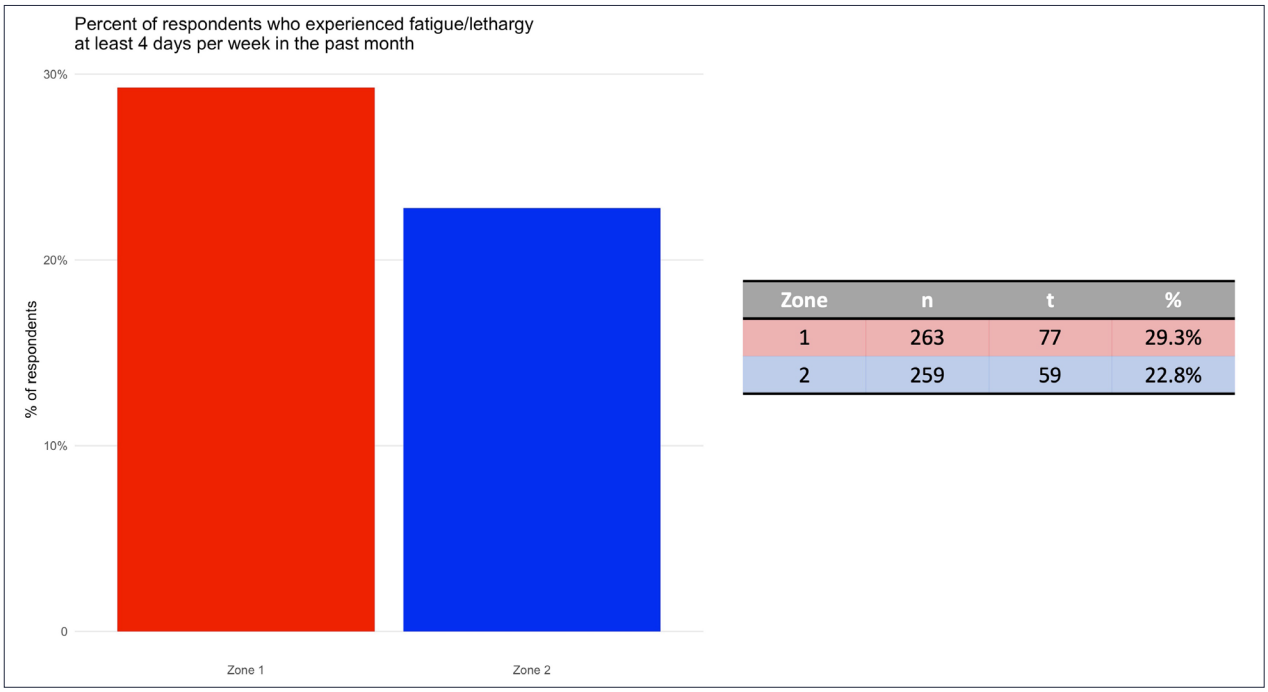


Zone	n	t	%
1	263	91	34.6%
2	259	79	30.5%

FIGURE 8.10 (LEFT)

Fatigue/lethargy

- **Nearly 30% of Zone 1 respondents experience fatigue/lethargy most of the time.**

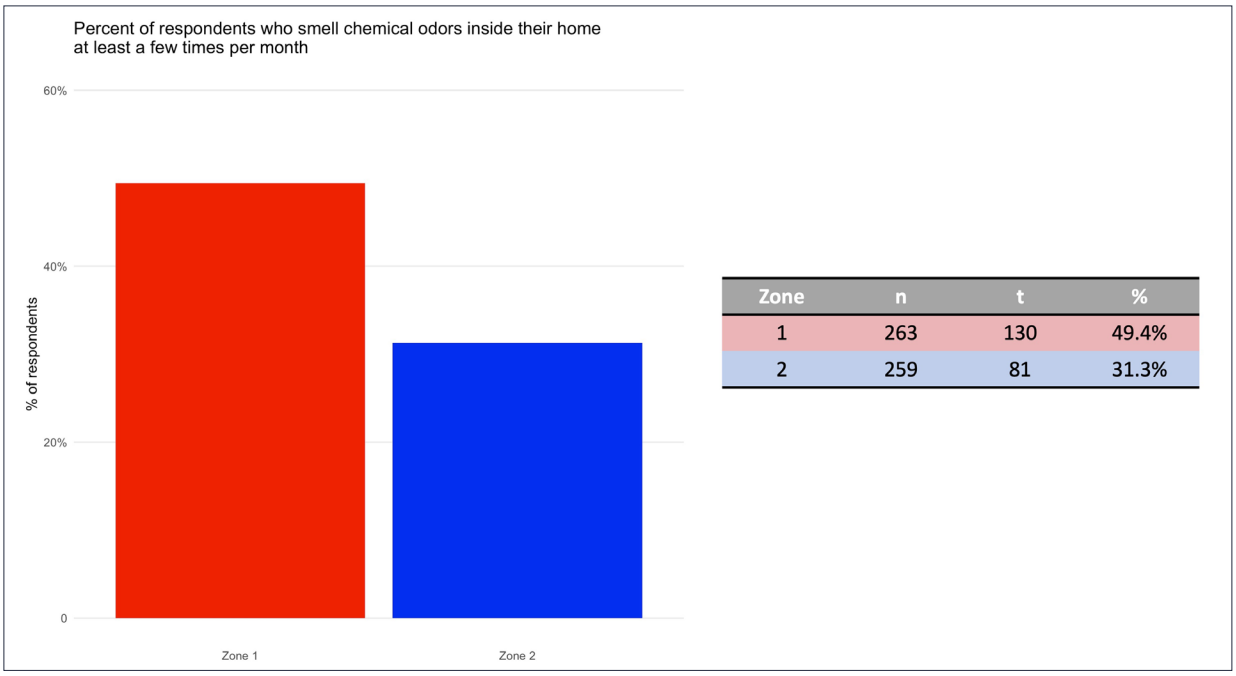


Zone	n	t	%
1	263	77	29.3%
2	259	59	22.8%

FIGURE 8.11 (LEFT)

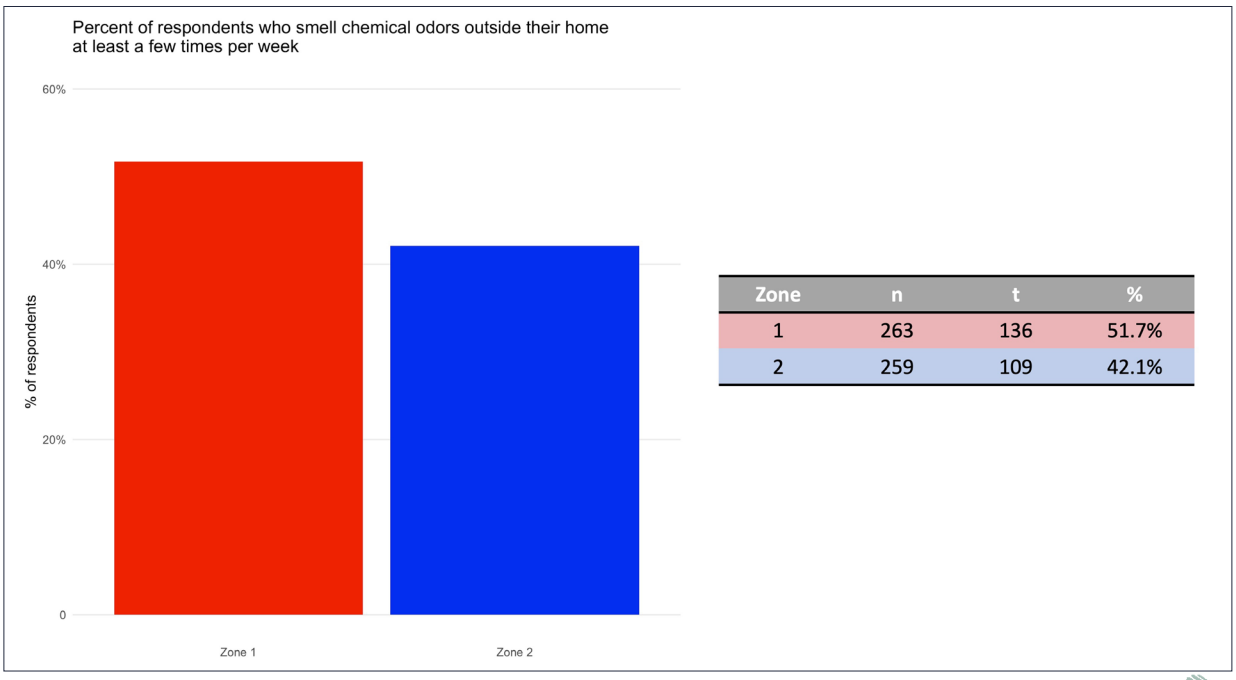
(IV) CHEMICAL ODORS

- **Nearly half of Zone 1 respondents smell chemical odors *inside* their homes at least a few times per month.**
- **Over half of Zone 1 respondents smell chemical odors outside their homes at least a few times per *week*.**
- **Over three-fourths of Zone 1 respondents smell chemical odors outside their homes at least a few times per month.**



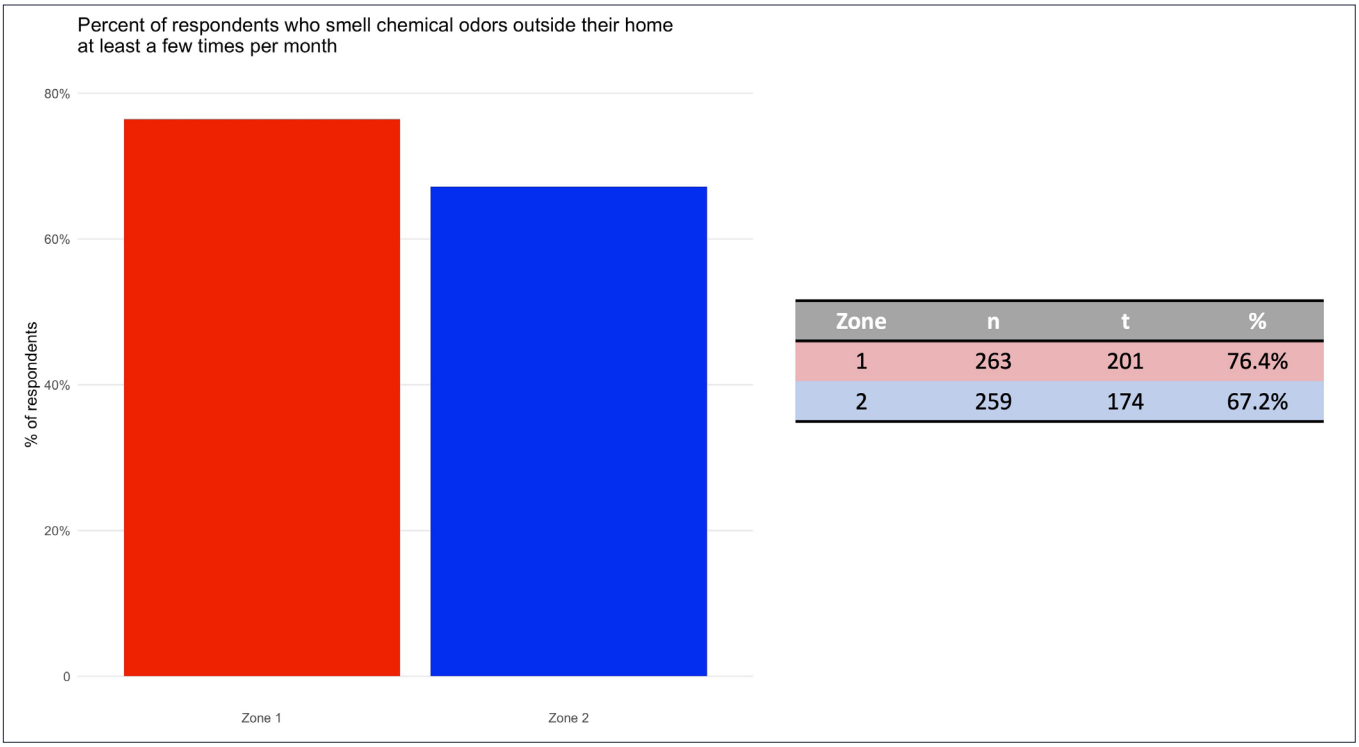
Zone	n	t	%
1	263	130	49.4%
2	259	81	31.3%

FIGURE 9.1 (LEFT)



Zone	n	t	%
1	263	136	51.7%
2	259	109	42.1%

FIGURE 9.2 (LEFT)

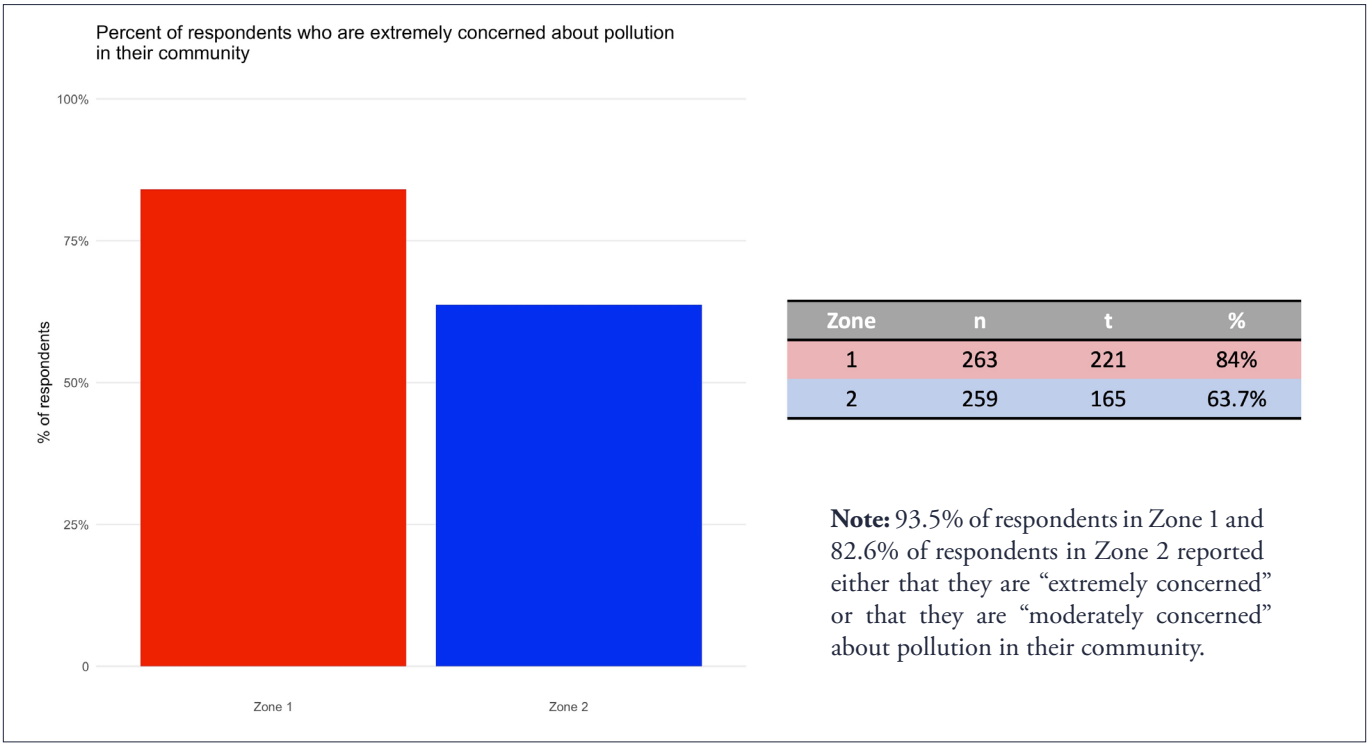


Zone	n	t	%
1	263	201	76.4%
2	259	174	67.2%

FIGURE 9.3 (ABOVE)

(V) CONCERN ABOUT POLLUTION

- **84% of Zone 1 respondents are "extremely concerned" about pollution in their community.**



Zone	n	t	%
1	263	221	84%
2	259	165	63.7%

Note: 93.5% of respondents in Zone 1 and 82.6% of respondents in Zone 2 reported either that they are "extremely concerned" or that they are "moderately concerned" about pollution in their community.

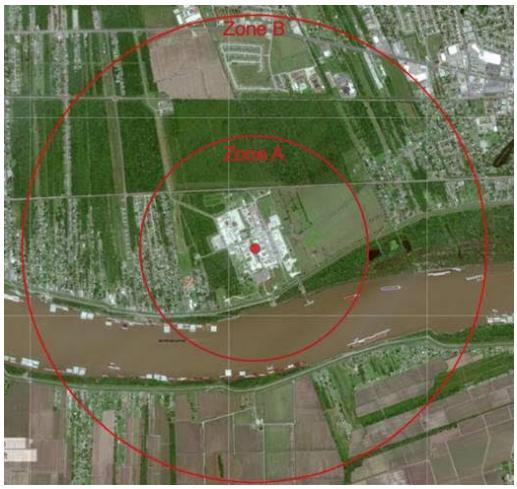
FIGURE 10.1 (ABOVE)

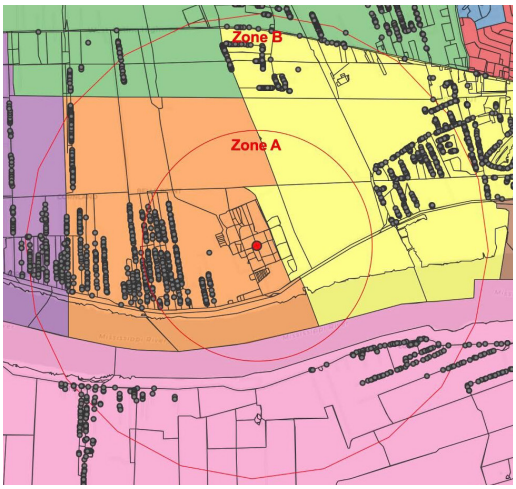
Appendix: Analyses by Spatial Zones A and B

Appendix: Analyses by Spatial Zones A and B

In addition to analyzing cancer prevalence and other health outcomes by spatial zones 1 and 2 (that is, the areas within 1.5 kilometers of the plant and between 1.5 and 2.5 kilometers from the plant, respectively), we also conducted analyses by a slightly different spatial zone breakdown. Below are the results of our analyses of the area within 1.25 kilometers of the plant ("Zone A") and the area between 1.25 and 2.5 kilometers from the plant ("Zone B").

Spatial zones A and B are illustrated in the images below. The outer circle circumscribes the entire survey area (the area within 2.5 kilometers of the plant) and the inner circle circumscribes the area within 1.25 kilometers of the plant. The plant—with a red dot at its center—can be seen at the center of the survey area.





Maps of the survey area. The outer circle circumscribes the area within a 2.5-km radius of the Denka facility. The inner circle circumscribes the area within a 1.25-km radius of the Denka facility. The facility can be seen at the center of the survey area. In the map on the right, the grey dots represent households and each color represents a different census tract. Residents of the orange-colored census tract face the highest risk in the country of developing cancer from air pollution, according to the EPA.

After eliminations (see *Data Set Used for Monte Carlo Analyses*, above), the total numbers of respondents, residents, and households included in the Monte Carlo analyses were as follows:

	Zone A	Zone B	Total
Households	188	337	525
Respondents	184	321	505
Residents	549	958	1,507

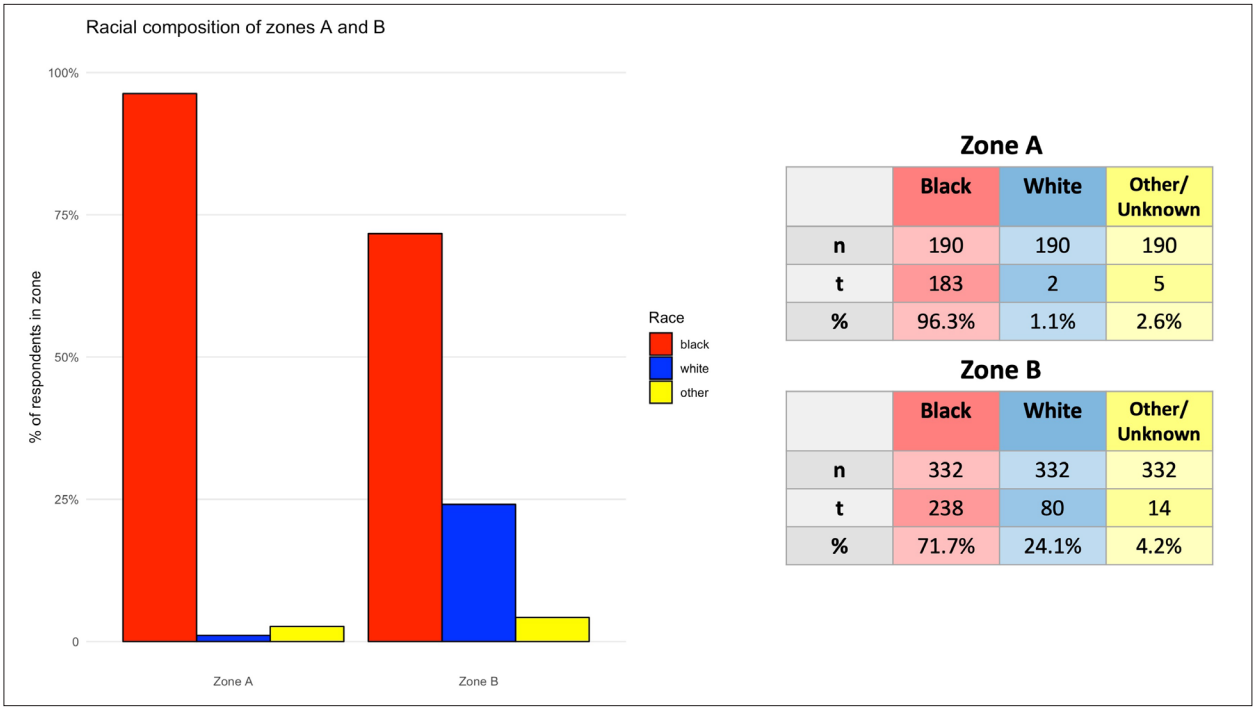
After eliminations (see *Introduction to Non-Cancer Health and Pollution Analyses*, above), the total numbers of respondents, residents, and households in the data set for our analyses of child health, symptoms, chemical odors, and concern about pollution were as follows:

	Zone A	Zone B	Total
Households	192	345	537
Respondents	190	332	522
Residents	559	984	1,543

After eliminations (see *Rapid pulse/rapid heart rate (tachycardia) diagnoses*, above), the total numbers of respondents and residents included in our Monte Carlo analyses of tachycardia prevalence were as follows:

	Zone A	Zone B	Total
Respondents	187	324	511
Residents	423	698	1,121

The following visualization provides a race/ethnicity breakdown of full-time respondents in Zones A and B:



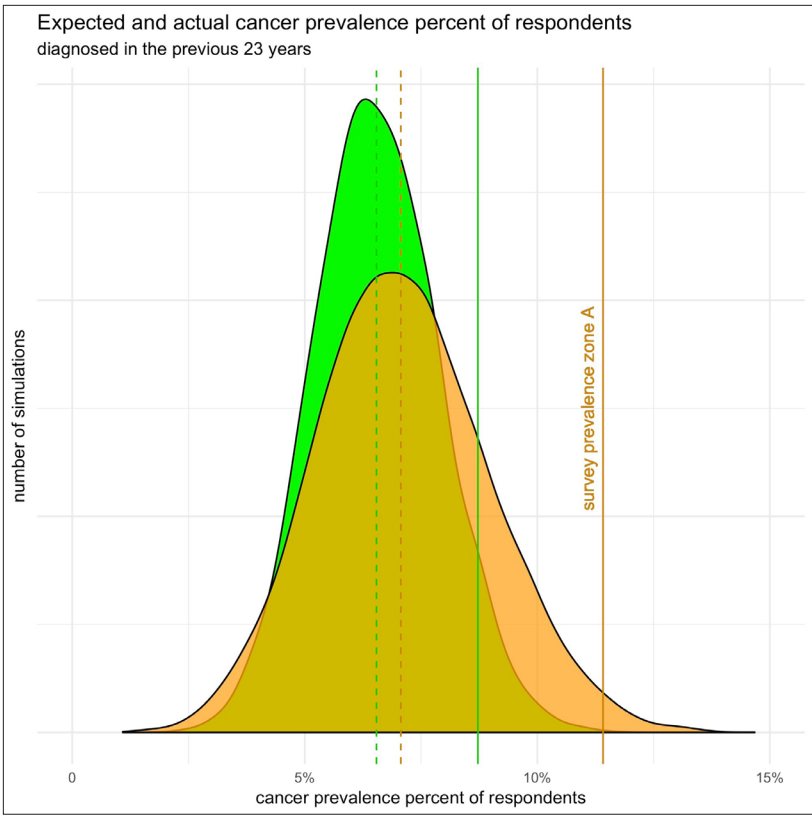
Zone A

	Black	White	Other/ Unknown
n	190	190	190
t	183	2	5
%	96.3%	1.1%	2.6%

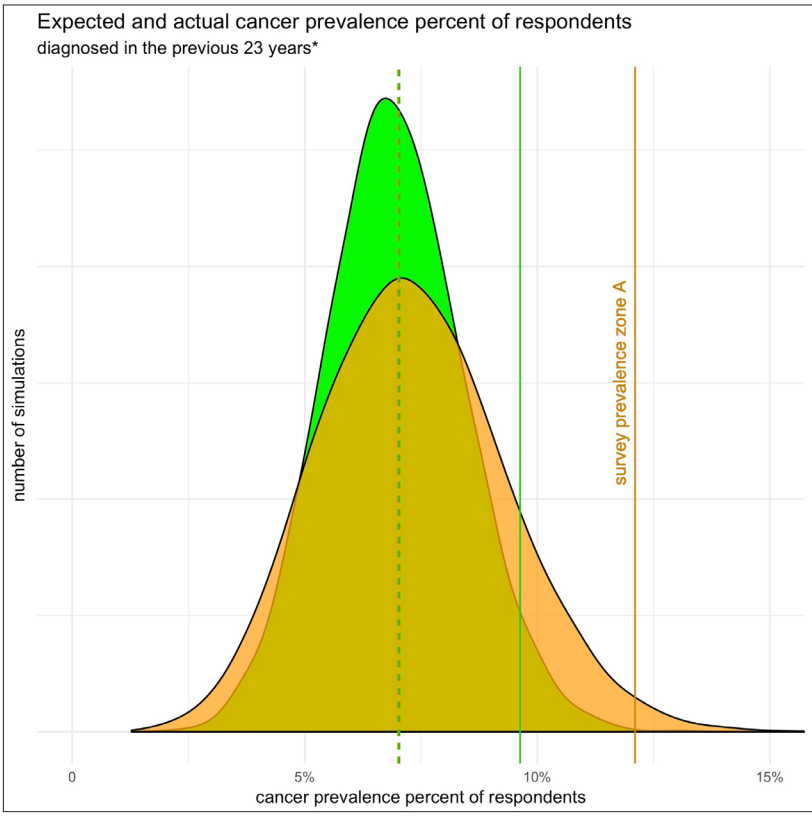
Zone B

	Black	White	Other/ Unknown
n	332	332	332
t	238	80	14
%	71.7%	24.1%	4.2%

Cancer prevalence among respondents, without and with smoking exclusion criterion:



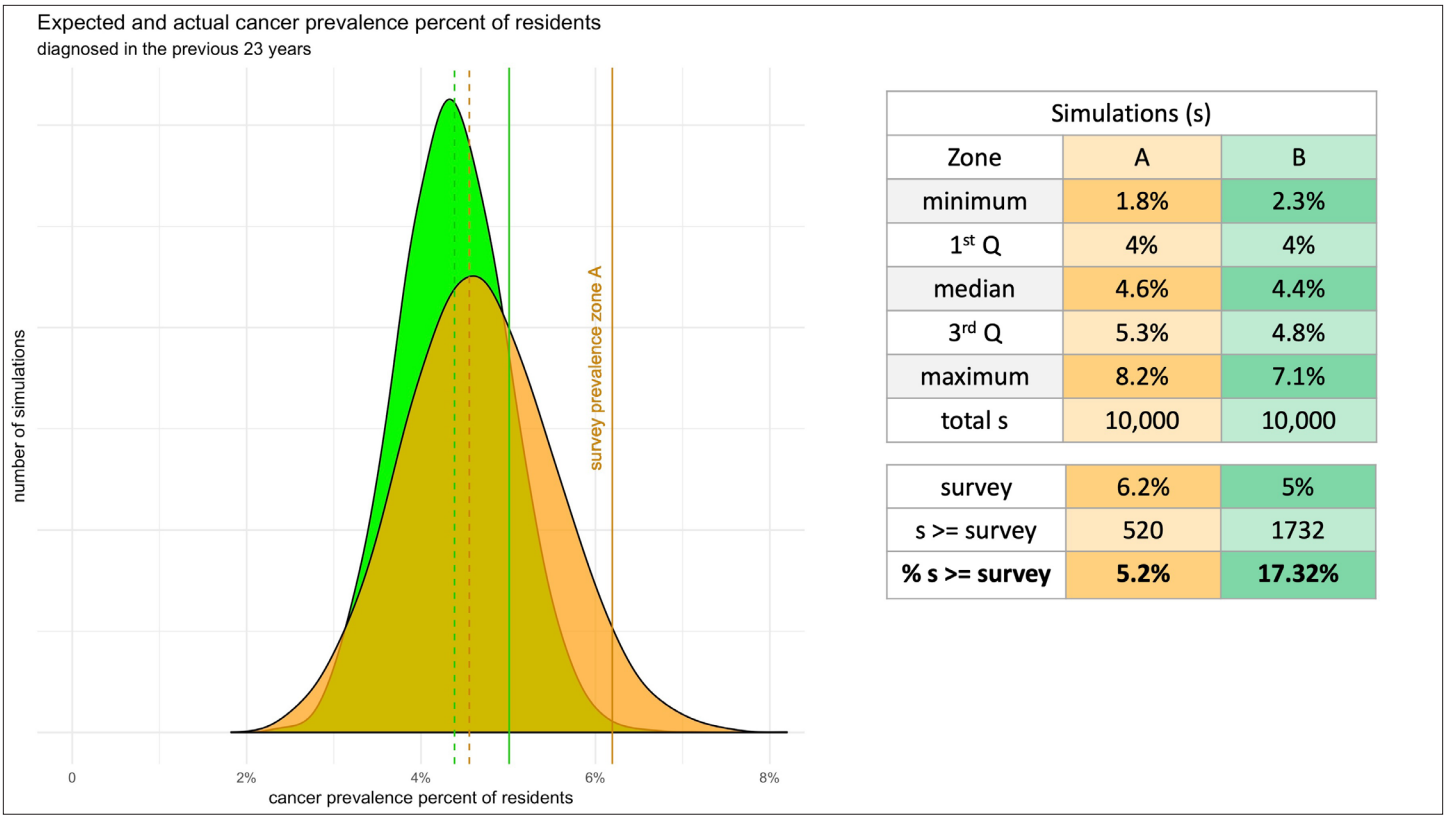
Simulations (s)		
Zone	A	B
minimum	1.1%	1.9%
1st Q	6%	5.6%
median	7.1%	6.5%
3rd Q	8.2%	7.5%
maximum	14.7%	13.4%
total s	10,000	10,000
survey	11.4%	8.7%
s >= survey	183	712
% s >= survey	**1.83%**	**7.12%**



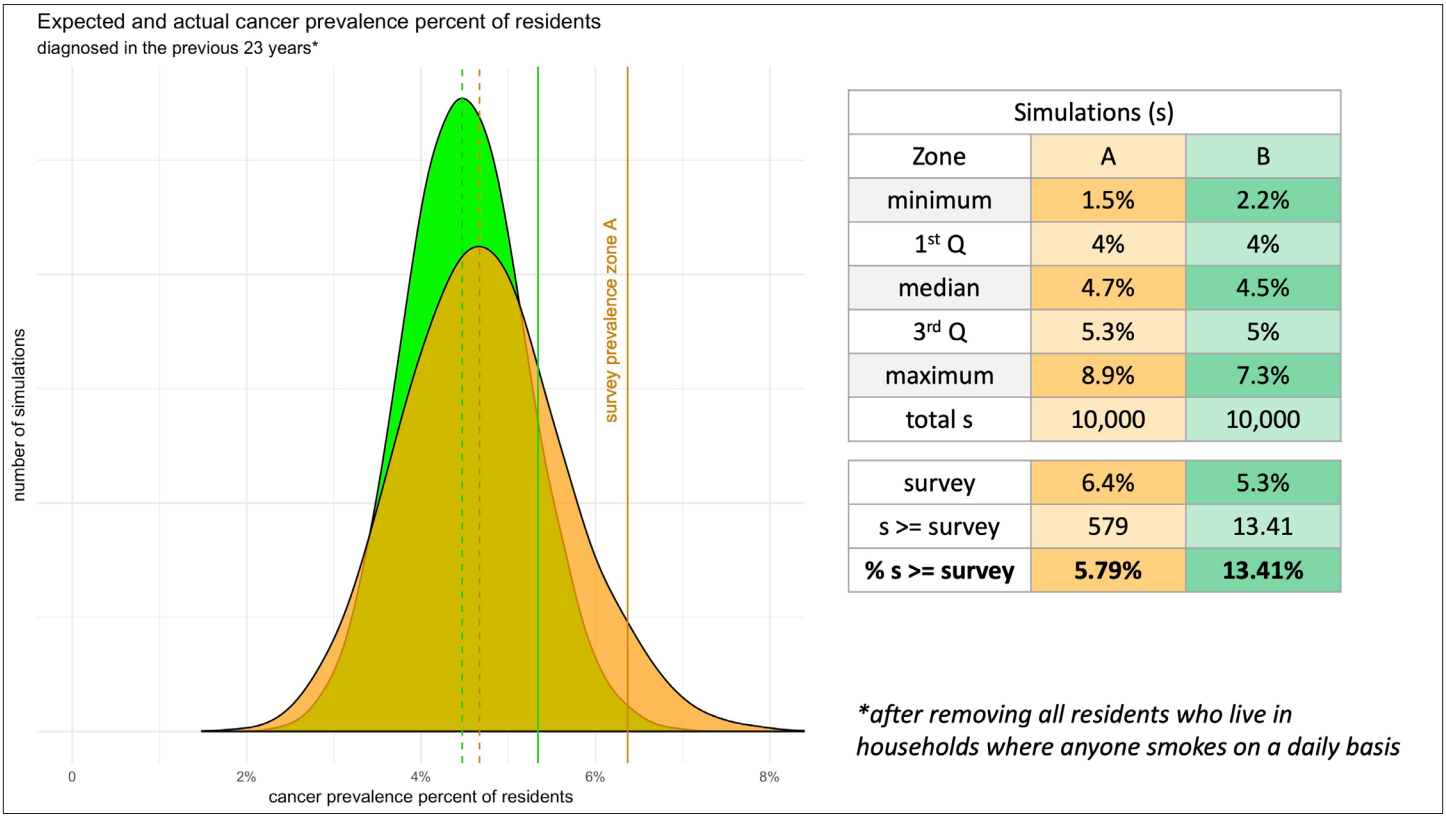
Simulations (s)		
Zone	A	B
minimum	1.3%	1.9%
1st Q	5.7%	5.9%
median	7%	7%
3rd Q	8.9%	7.8%
maximum	15.9%	13.3%
total s	10,000	10,000
survey	12.1%	9.6%
s >= survey	177	477
% s >= survey	**1.77%**	**4.77%**

**after removing all respondents who live in households where anyone smokes on a daily basis*

Cancer prevalence among residents, without and with smoking exclusion criterion:



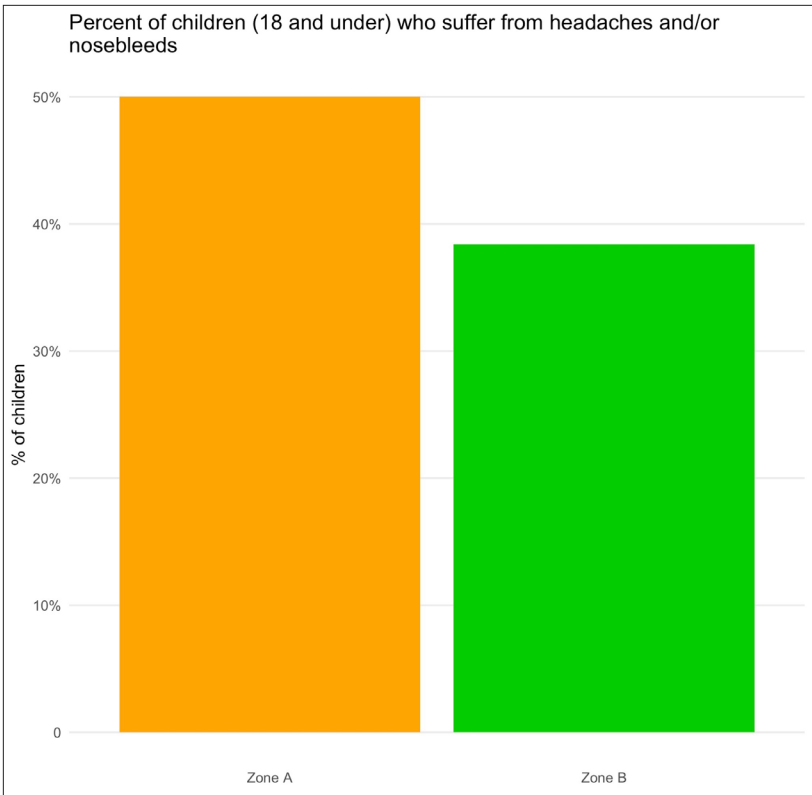
Simulations (s)		
Zone	A	B
minimum	1.8%	2.3%
1st Q	4%	4%
median	4.6%	4.4%
3rd Q	5.3%	4.8%
maximum	8.2%	7.1%
total s	10,000	10,000
survey	6.2%	5%
s >= survey	520	1732
% s >= survey	**5.2%**	**17.32%**



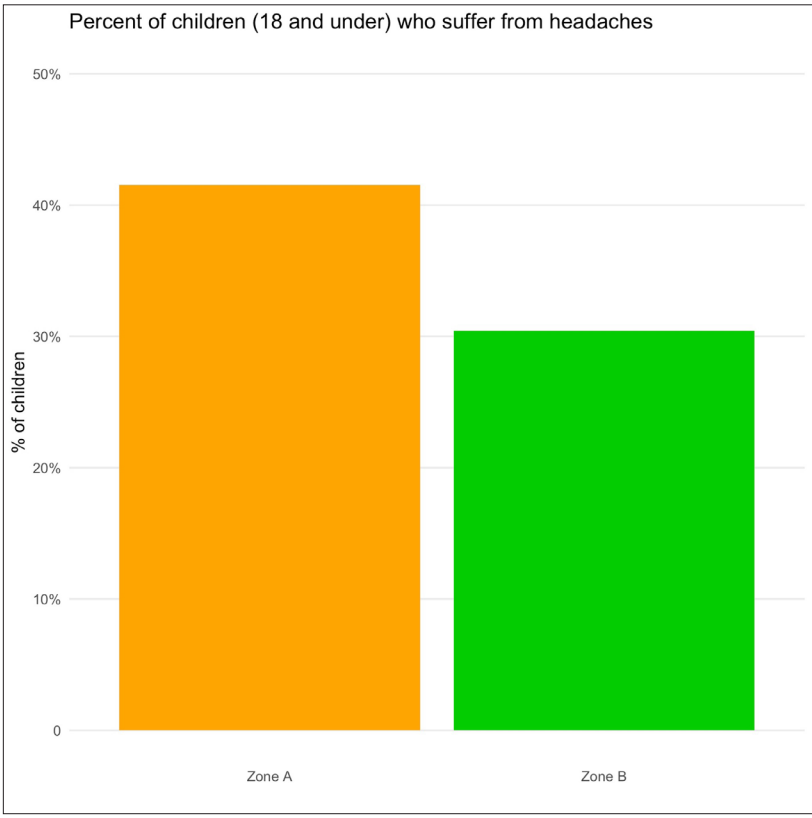
Simulations (s)		
Zone	A	B
minimum	1.5%	2.2%
1st Q	4%	4%
median	4.7%	4.5%
3rd Q	5.3%	5%
maximum	8.9%	7.3%
total s	10,000	10,000
survey	6.4%	5.3%
s >= survey	579	13.41
% s >= survey	**5.79%**	**13.41%**

**after removing all residents who live in households where anyone smokes on a daily basis*

Child health:



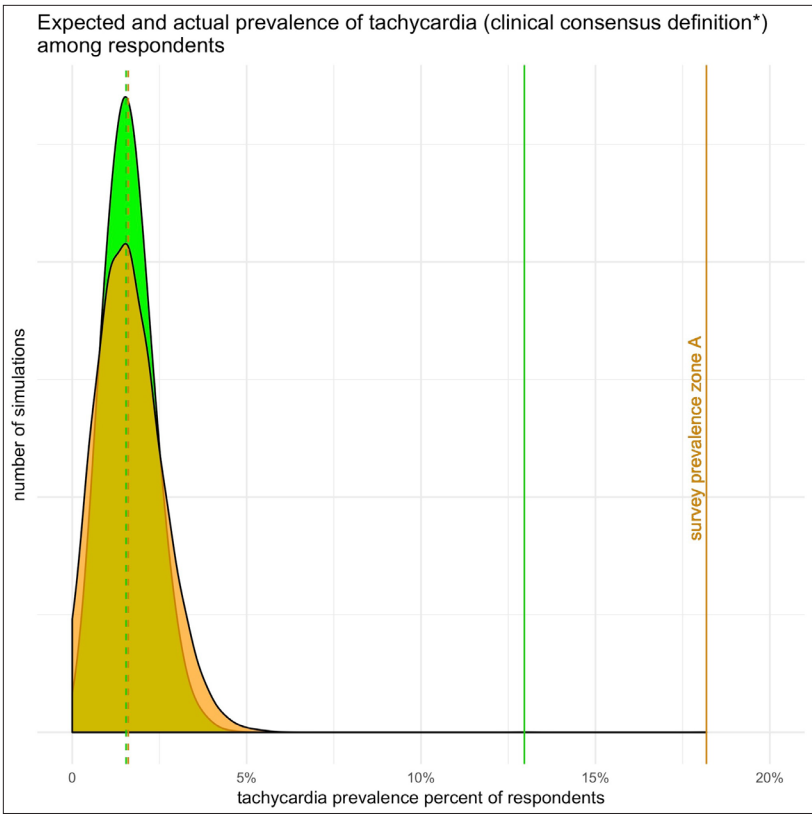
Zone	n	t	%
A	130	65	50%
B	276	106	38.4%



Zone	n	t	%
A	130	54	41.5%
B	276	84	30.4%

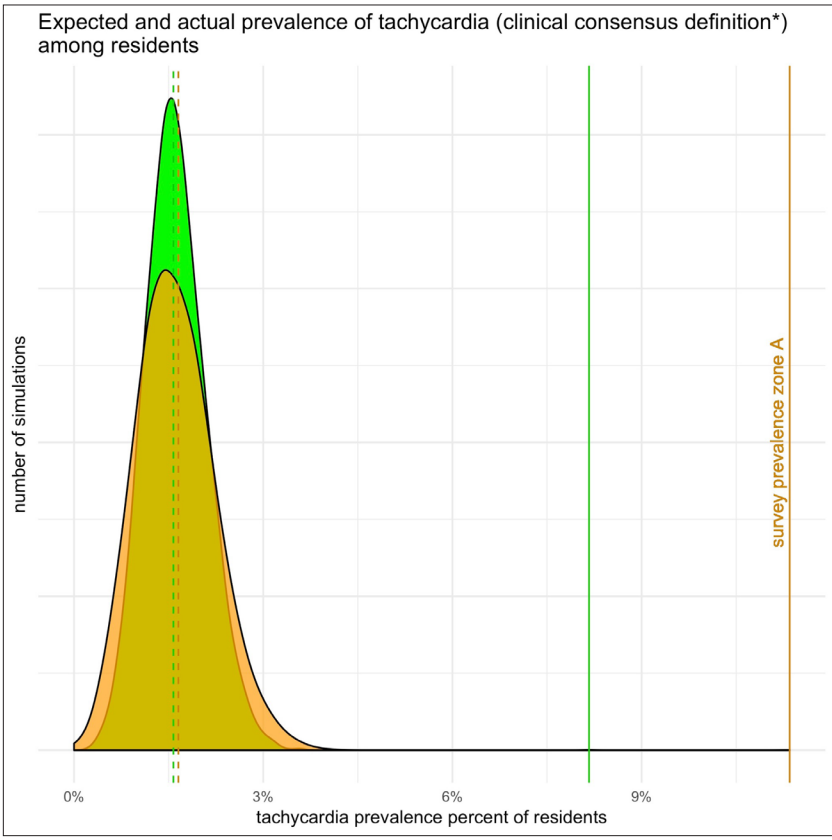
Zone	n	t	%
A	130	28	21.5%
B	276	51	18.5%

Rapid pulse/rapid heart rate (tachycardia) diagnoses:



Simulations (s)		
Zone	A	B
minimum	0%	0%
1st Q	1.1%	1.2%
median	1.6%	1.5%
3rd Q	2.1%	2.2%
maximum	5.9%	4.9%
total s	10,000	10,000
survey	**18.2%**	**13%**
s >= survey	0	0
% s >= survey	**0%**	**0%**

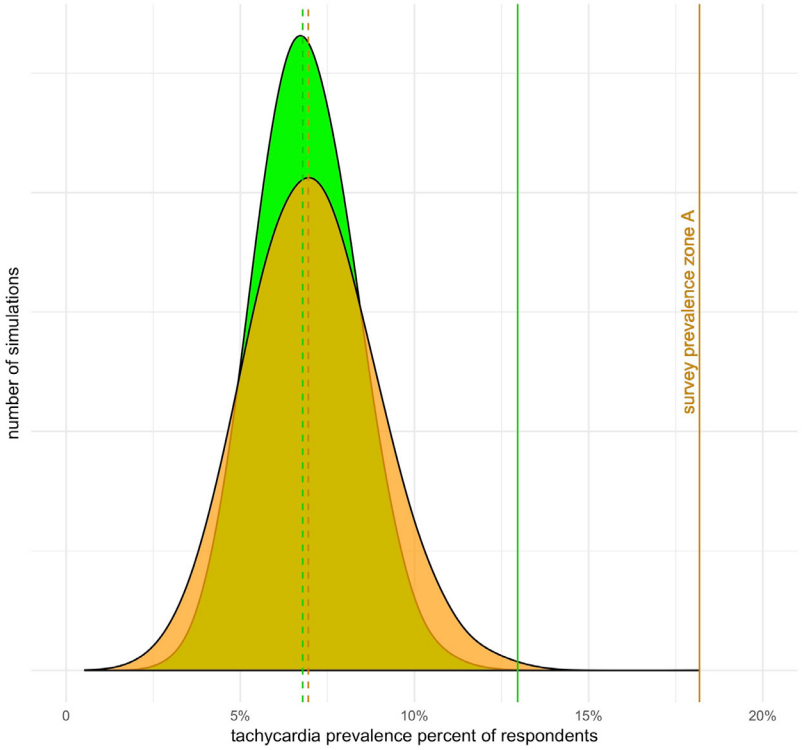
*resting pulse rate (RPR) greater than 100 beats/minute



Simulations (s)		
Zone	A	B
minimum	0%	0.3%
1st Q	1.2%	1.3%
median	1.7%	1.6%
3rd Q	1.9%	1.9%
maximum	4.3%	4.2%
total s	10,000	10,000
survey	**11.3%**	**8.2%**
s >= survey	0	0
% s >= survey	**0%**	**0%**

*resting pulse rate (RPR) greater than 100 beats/minute

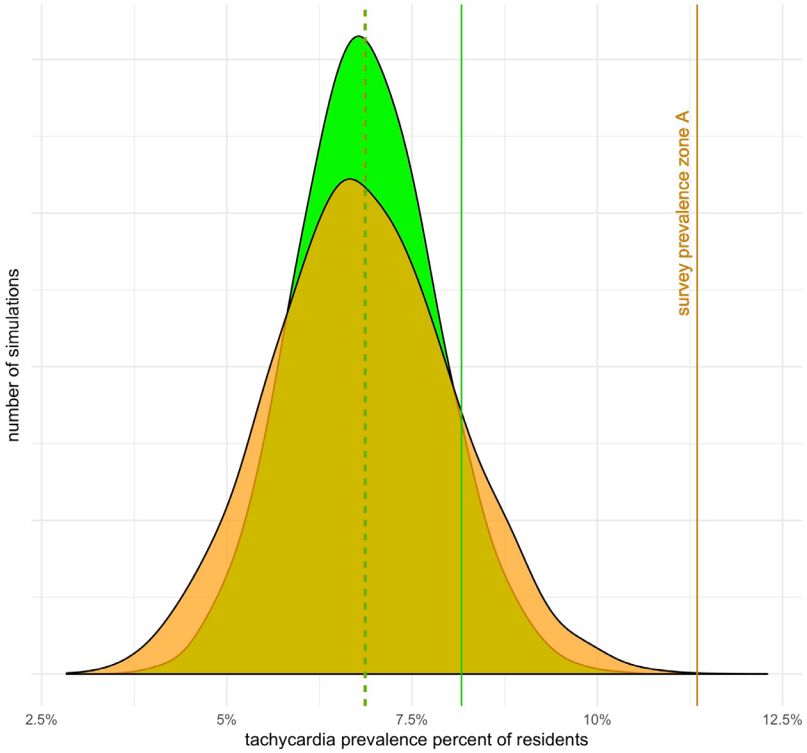
Expected and actual prevalence of tachycardia (revised clinical guideline*) among respondents



Simulations (s)		
Zone	A	B
minimum	0.5%	2.8%
1st Q	5.9%	5.9%
median	7%	6.8%
3rd Q	8%	7.7%
maximum	14.3%	12.3%
total s	10,000	10,000
survey	**18.2%**	**13%**
s >= survey	0	0
% s >= survey	**0%**	**0%**

*resting pulse rate (RPR) greater than 90 beats/minute

Expected and actual prevalence of tachycardia (revised clinical guideline*) among residents

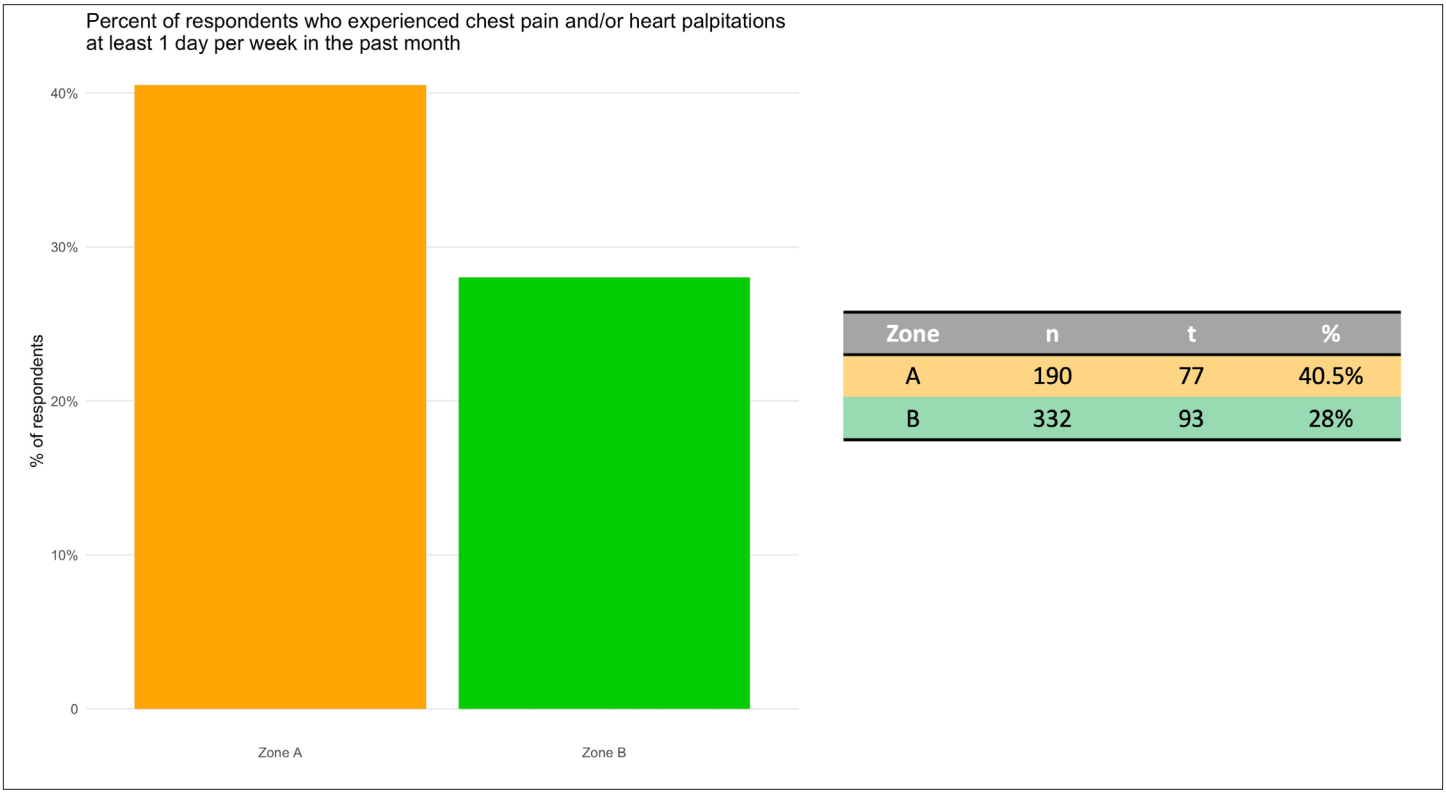


Simulations (s)		
Zone	A	B
minimum	2.8%	3.6%
1st Q	5.9%	6.2%
median	6.9%	6.9%
3rd Q	7.6%	7.4%
maximum	12.3%	11.2%
total s	10,000	10,000
survey	**11.3%**	**8.2%**
s >= survey	3	1008
% s >= survey	**0.03%**	**10.08%**

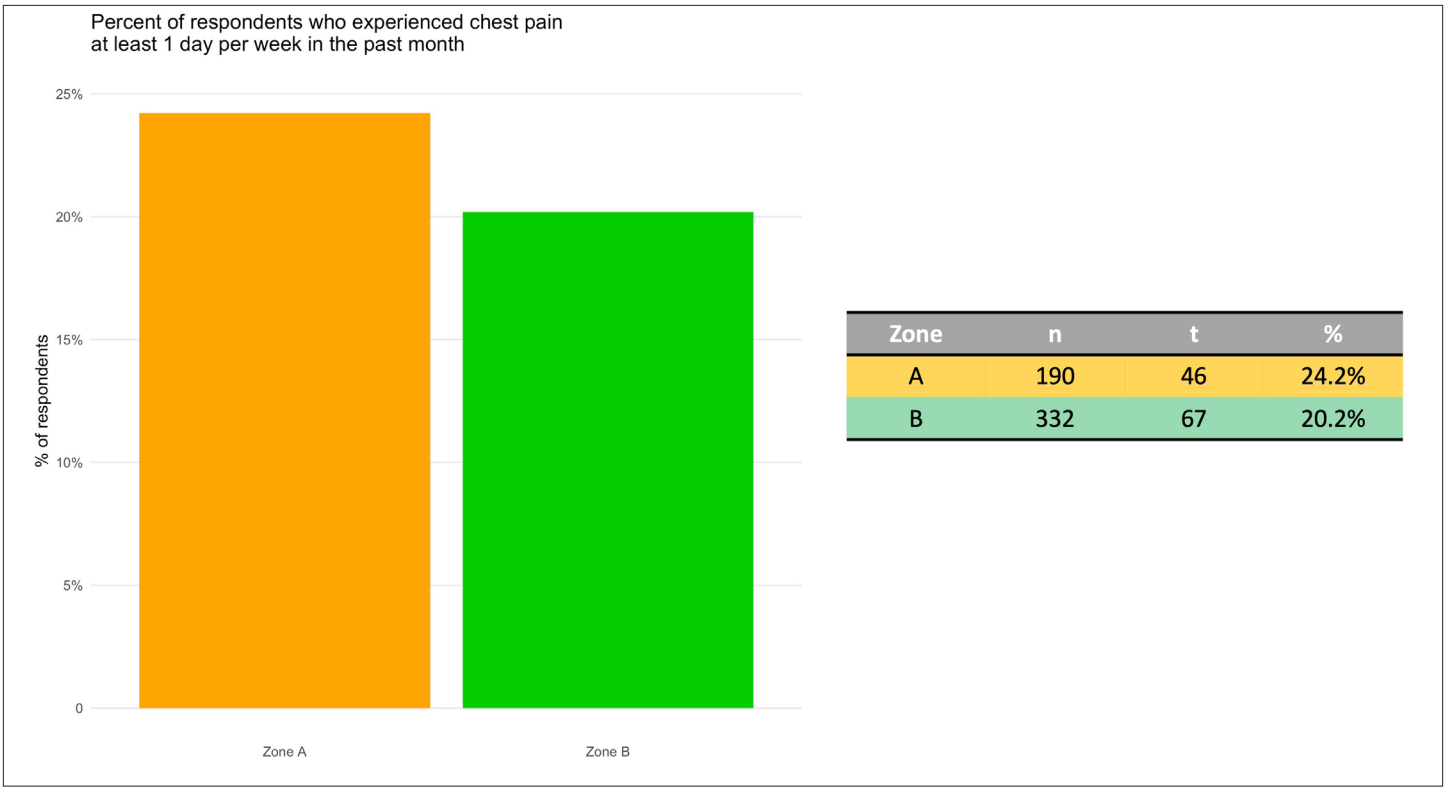
*resting pulse rate (RPR) greater than 90 beats/minute

Symptoms:

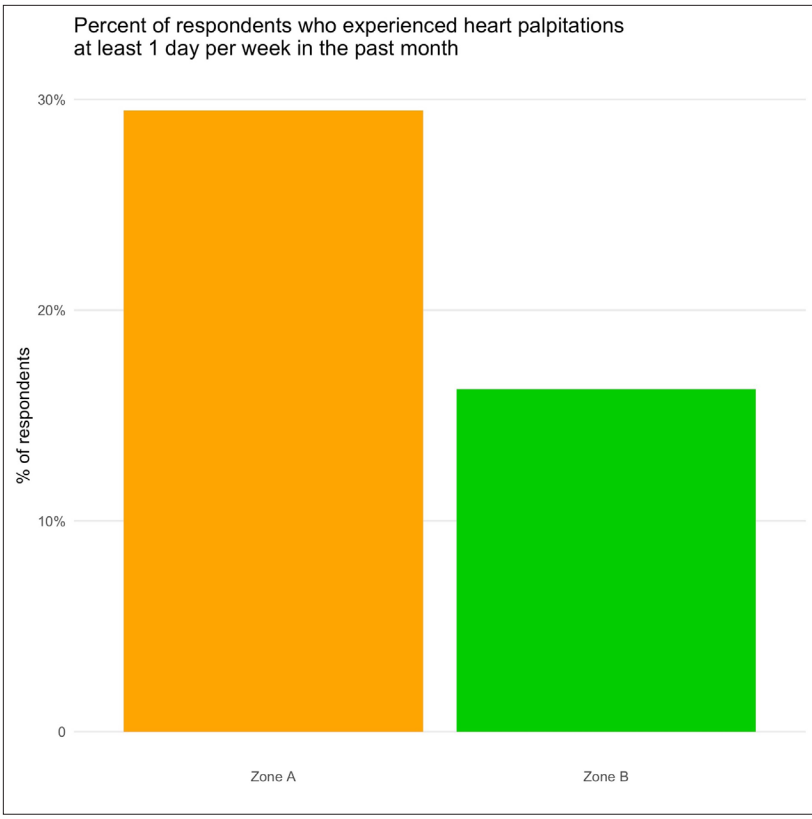
Chest pain and heart palpitations



Zone	n	t	%
A	190	77	40.5%
B	332	93	28%

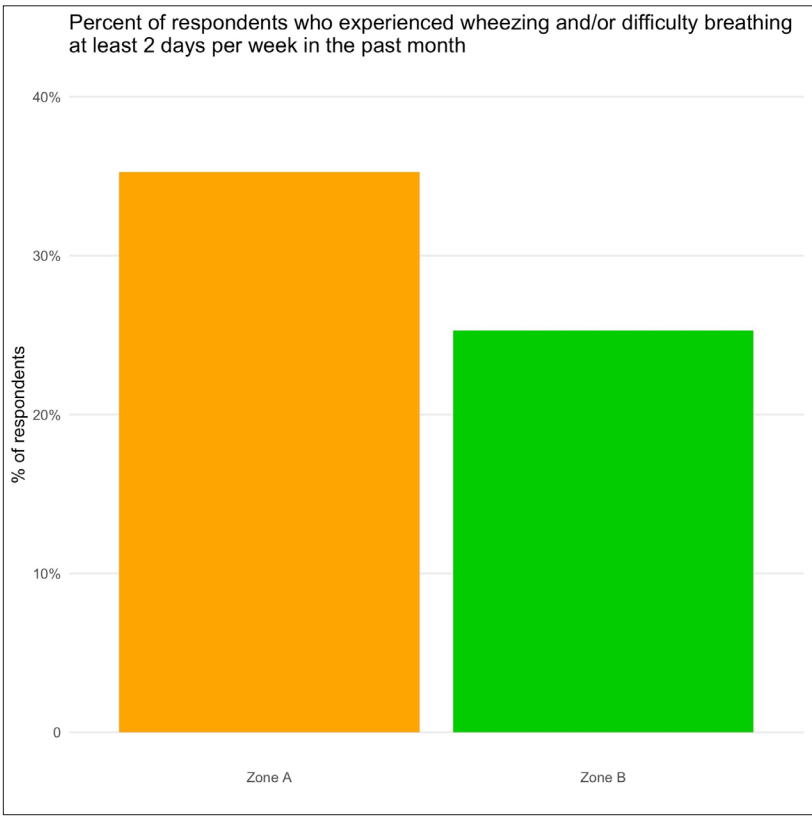


Zone	n	t	%
A	190	46	24.2%
B	332	67	20.2%



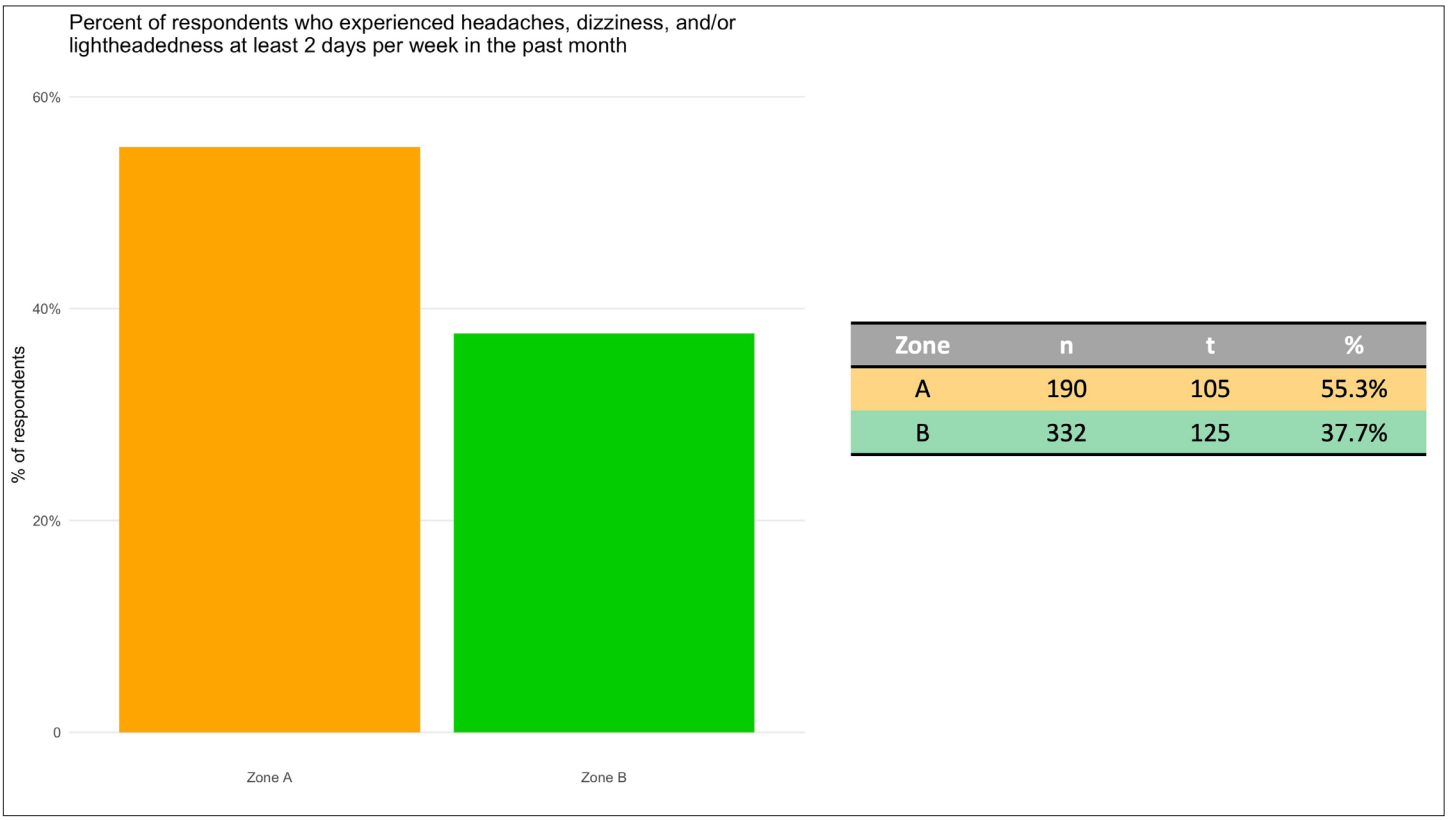
Zone	n	t	%
A	190	56	29.5%
B	332	54	16.3%

Wheezing and difficulty breathing

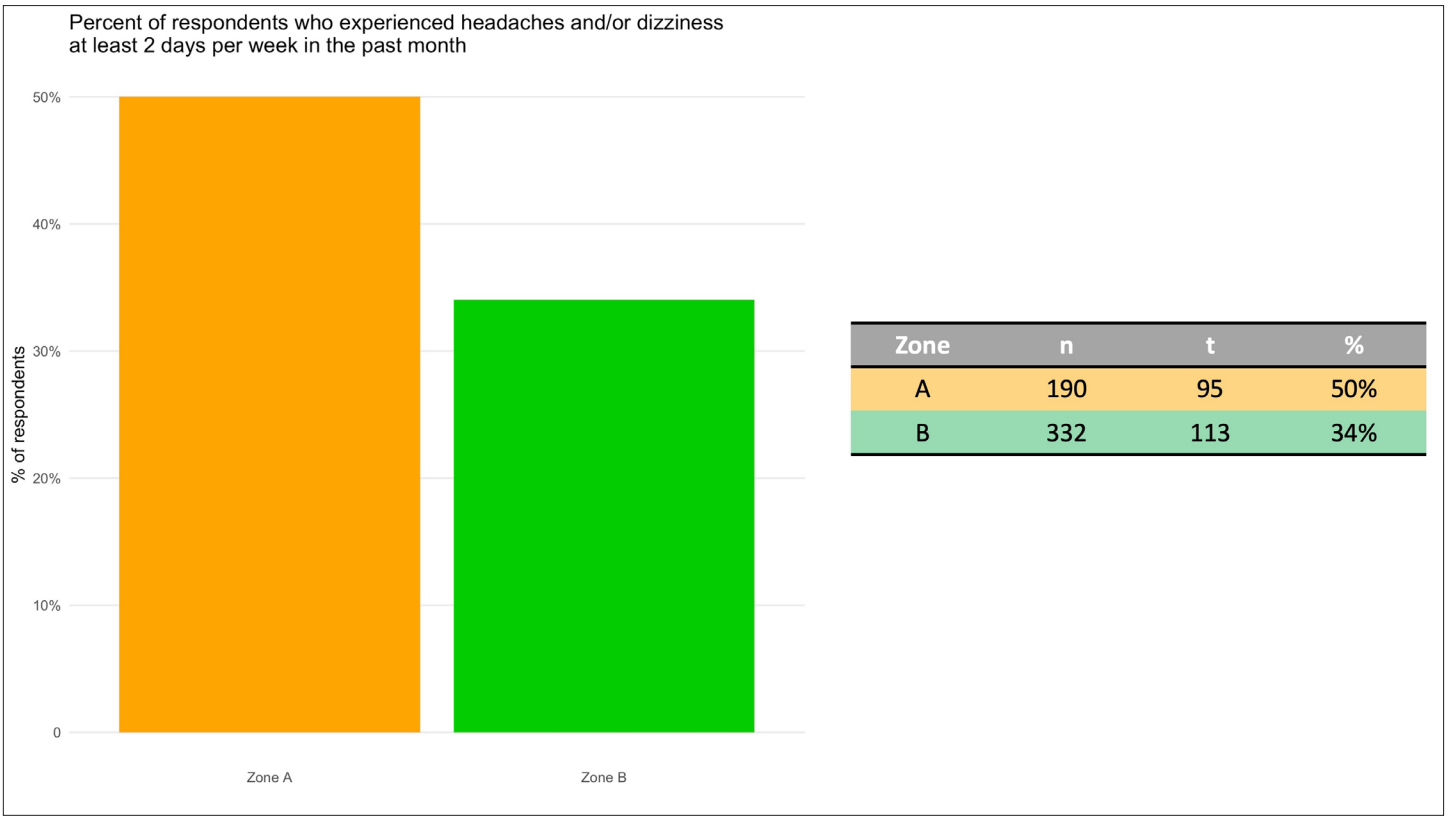


Zone	n	t	%
A	190	67	35.3%
B	332	84	25.3%

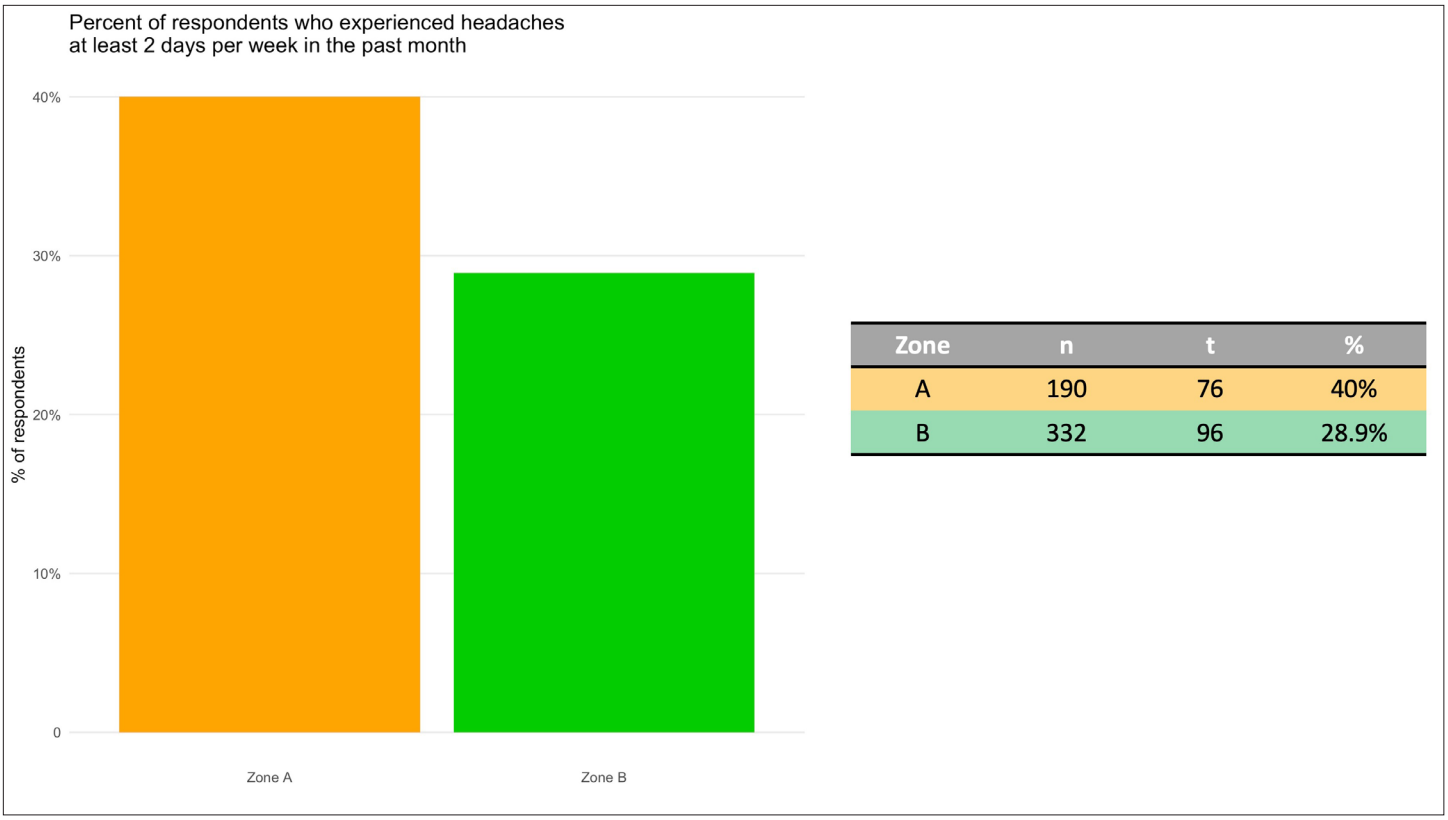
Headaches, dizziness, and lightheadedness



Zone	n	t	%
A	190	105	55.3%
B	332	125	37.7%

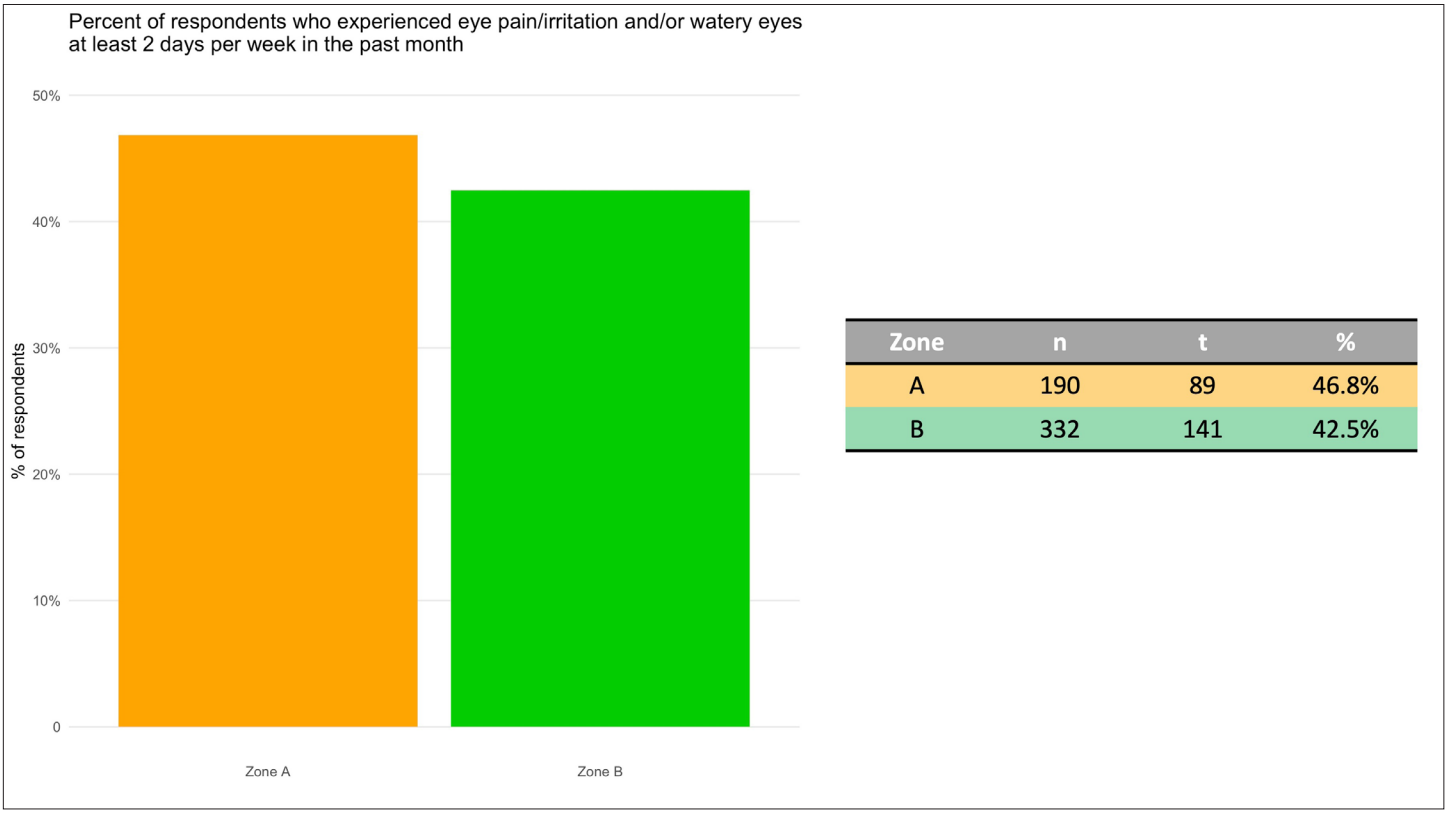


Zone	n	t	%
A	190	95	50%
B	332	113	34%



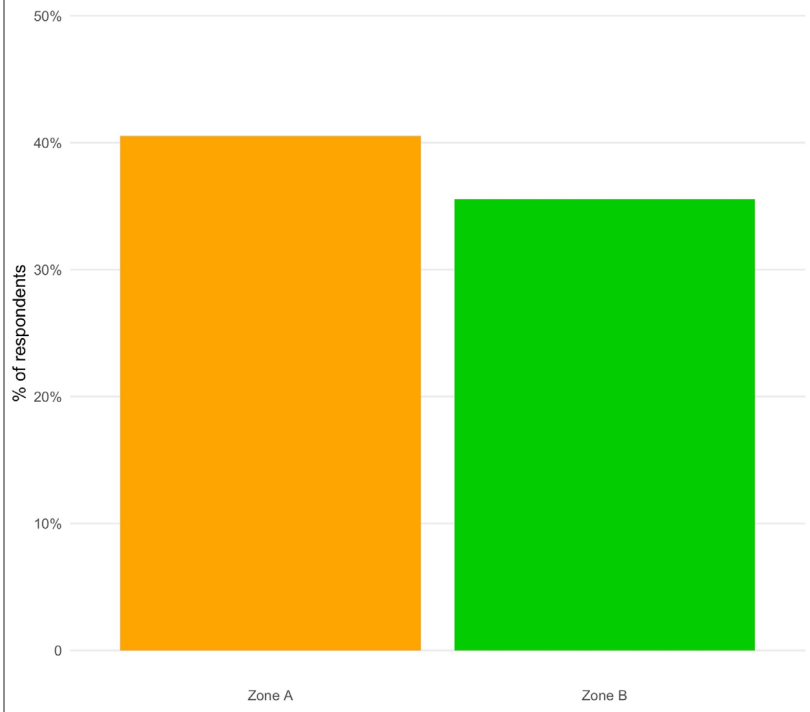
Zone	n	t	%
A	190	76	40%
B	332	96	28.9%

Eye pain/irritation and watery eyes



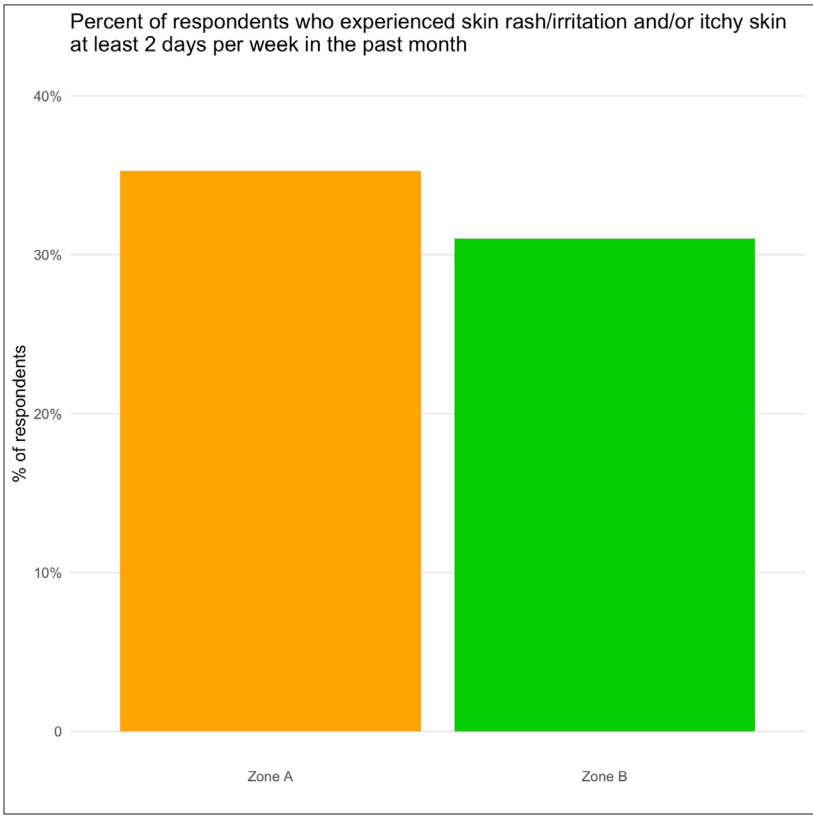
Zone	n	t	%
A	190	89	46.8%
B	332	141	42.5%

Cough, sneezing, and sore/hoarse throat



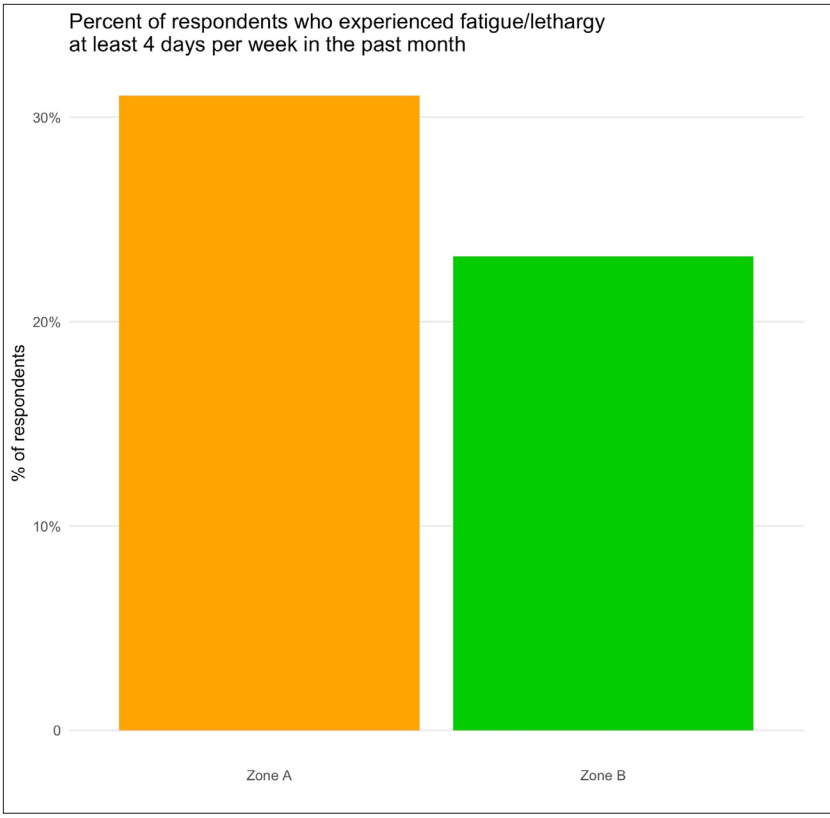
Zone	n	t	%
A	190	77	40.5%
B	332	118	35.5%

Skin rash/irritation and itchy skin



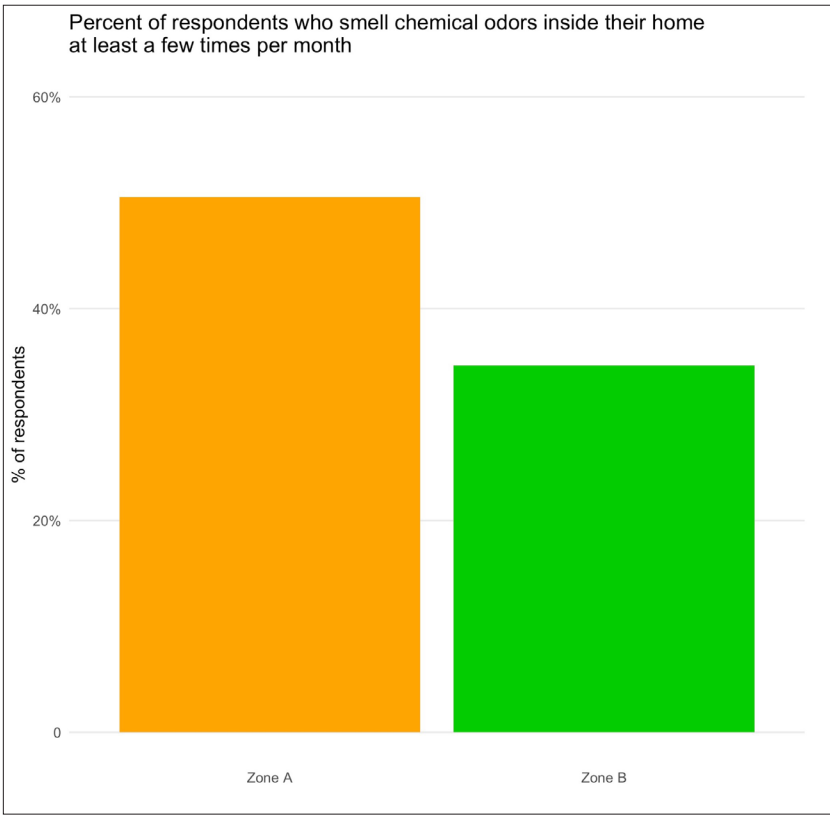
Zone	n	t	%
A	190	67	35.3%
B	332	103	31%

Fatigue/lethargy



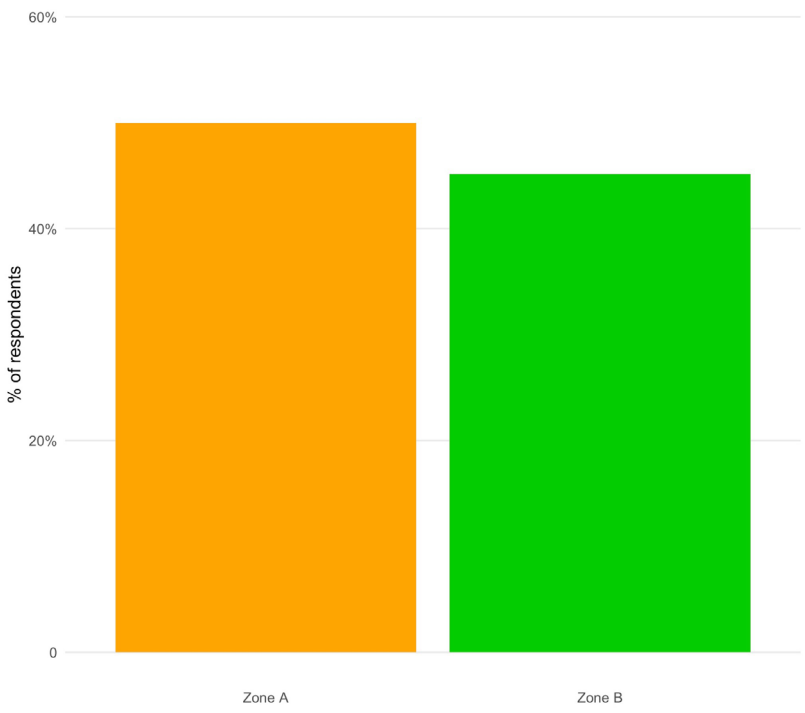
Zone	n	t	%
A	190	59	31.1%
B	332	77	23.2%

Chemical odors:



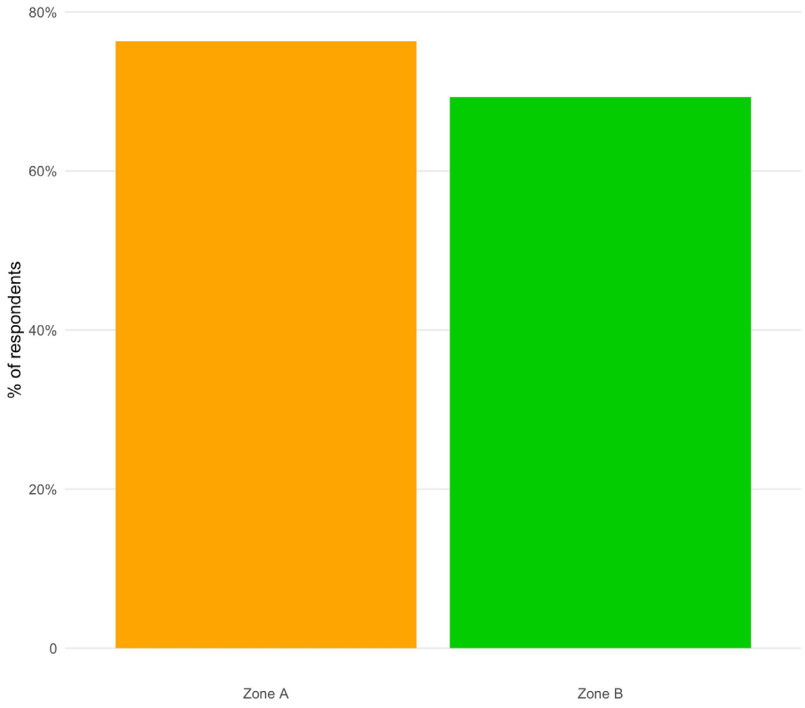
Zone	n	t	%
A	190	96	50.5%
B	332	115	34.6%

Percent of respondents who smell chemical odors outside their home at least a few times per week



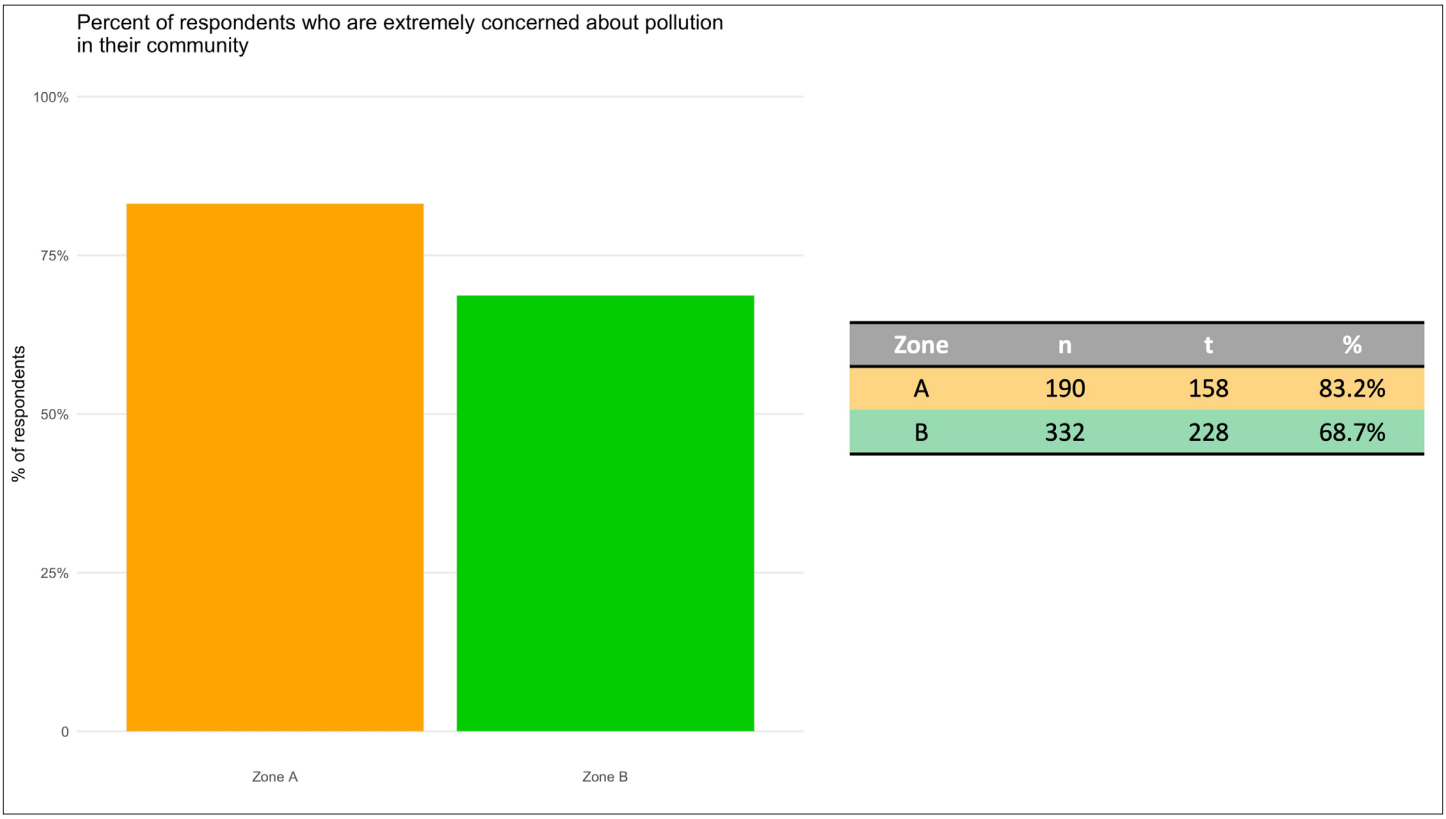
Zone	n	t	%
A	190	95	50%
B	332	150	45.2%

Percent of respondents who smell chemical odors outside their home at least a few times per month



Zone	n	t	%
A	190	145	76.3%
B	332	230	69.3%

Concern about pollution:



Zone	n	t	%
A	190	158	83.2%
B	332	228	68.7%

UNIVERSITY NETWORK FOR HUMAN RIGHTS