LAKEHEAD UNIVERSITY

A STUDY OF THE RELATIONSHIP BETWEEN HEALTH SYMPTOM REPORTING AND MOLD MEASUREMENTS IN FIRST NATIONS HOUSING

PROJECT SUBMITTED TO
THE FACULTY OF PROFESSIONAL SCHOOLS
IN CANDIDACY FOR THE DEGREE OF MASTER OF PUBLIC HEALTH

BY
ERIK ALLAIN

THUNDER BAY, ONTARIO
OCTOBER 24, 2008
STUDENT ID: 0318289

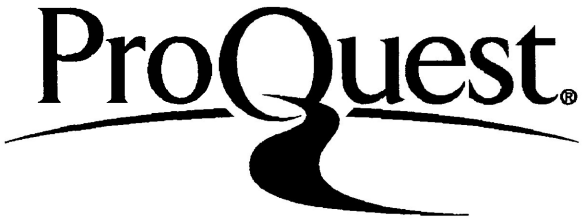

TABLE OF CONTENT

ACKNOWLEDGMENTS

I offer my sincerest gratitude to my supervisor, Dr William Montelpare, who has supported me throughout this project with his patience, and knowledge whilst allowing me the room to work in my own way. Dr Montelpare was abundantly helpful and offered invaluable assistance, support and guidance.

1. Introduction

The purpose of the study is to examine the relationship between health symptom reporting and mold concentration in First Nations housing. Indoor air quality is increasingly becoming a topic of concern for many homeowners, and the information on the issue is often controversial and misunderstood. Various studies report that a range of health symptoms are associated with being exposed to mold in the indoor environment, but there is still debate over which symptoms are included in this association and the amount of mold to which an individual must be exposed before the symptoms occur.

One study concludes that, while there is general agreement that visible mold in the indoor environment is unhealthy and is a situation that must be corrected, the point at which mold contamination becomes a health concern is not known (Robbins et al., 2000). Other studies state that people experiencing mold-related health effects have diverse characteristics, and that, as there are currently no health-based exposure standards for molds, it is difficult, if not impossible, to predict who will experience symptoms and when (Bobbit et al., 2005; Rao et al., 1996). Even though these studies have led to a significant amount of scientific literature, unanimous agreement has not been reached as yet on whether or not mold in fact causes health symptoms in the individuals exposed.

This subject is important, considering that it is reported that Canadians spend up to 70% of their time indoors—perhaps even more in the coldest months of the year (1a). Therefore, when houses are contaminated with mold, the exposure time of the occupants is significant and ongoing. Housing inequalities among individuals living on and off reserve have also been recognized for some time. According to a report produced by the National Aboriginal Health Organization (RHS, 2002/03), nearly half (49%) of families living in band-owned houses report mold in their homes. This was confirmed in the *2003 Report of the Auditor General of Canada* (RAGC, 2003), which stated: "Mold contaminates almost half of First Nations households."

First Nations homes are also about four times more likely to require extensive repairs than are Canadian homes overall (34% vs. 8%) (RHS, 2002/03). Exacerbating the problem is the fact that First Nations members live in homes that are, on average, nearly twice as crowded as Canadian homes overall (4.8 vs. 2.6 persons per room) (RHS, 2002/03).

There are significant gaps in the literature regarding the health effects of mold on First Nations individuals. Very few studies have been conducted in which First Nations people were the target population and in which mold measurements and reported health symptoms were collected in a systematic fashion. This study was designed to bridge this gap by

examining the association between reported health symptoms and mold concentration in an entire First Nations community.

2. Project Background

The study was conducted in the First Nations community of Neskantaga. Neskantaga is a remote First Nations reserve in the northern reaches of the Canadian province of Ontario, in the District of Kenora (52.191402° N, 88.03638° W). The total population of Neskantaga according to the 2006 Statistics Canada census consisted of 260 individual divided between 140 males and 125 females. The median age of the population was 21.7 years of age, of which 61.5% were aged 15 years and older. Age characteristics of the population can be found in table x. During this study, 193 individuals answered the questionnaire for a response rate of 73%. Only one individual refused to answer the questionnaire. The other individuals we couldn't reach were outside of the community during the study period.

A total of 74 private dwellings were counted in the community. We were able to inspect 59 of these private dwellings therefore we were given access to 80% of all the private dwellings. Of the 15 houses we were not able to inspect, in one case the owner did not want us to inspect the dwelling, in another case the house damaged by fire and was unoccupied, three houses were condemned and thus unoccupied as well. The other ten houses were occupied but the owners were out of town during our time in the community. On average, there were 3.6 individuals per dwelling.

Table 1: Population and dwelling counts for Neskantaga First Nation

Population and dwelling counts	Neskantaga First Nation total	Male	Female
Population 2006	260	140	125
Number of private dwellings	74		
Average of people per dwelling	3.6 per dwelling		
Number of dwellings inspected	59 dwellings		
% of total dwellings inspected	80%		
Number of individuals who answered questionnaire	193		
Response rate	73%		

Table 2: Age characteristics of the Aboriginal population

Age characteristics of the Aboriginal population	Neskantaga First Nation total	Male	Female
Population 2006	260	140	125
0 to 4	40	25	15
5 to 9	25	10	15
10 to 14	35	15	20
15 to 19	20	15	0
20 to 24	25	15	15
25 to 29	20	10	10
30 to 34	15	10	15
35 to 39	15	10	0
40 to 44	15	10	10
45 to 49	10	0	10
50 to 54	10	10	0
55 to 59	10	0	10
60 to 64	10	10	0
65 to 69	0	10	0
70 to 74	10	0	10
75 to 79	0	0	0
80 to 84	0	0	0
85 years and over	0	0	0
Median age of the Aboriginal identity population	21.7	21.5	21.7
% of the Aboriginal identity population aged 15 and over	61.5	60.7	60.0

Many houses in this community had been reported to be contaminated with varying levels of visible indoor mold. The community health board and its representatives had investigated several homeowners' complaints of mold contamination and had reported that they believed many homes were indeed contaminated. However, no comprehensive, community-wide mold survey had ever been conducted in this community. Many community residents had also complained of mold-exposure-related health effects, but, in many cases, it was not known whether their houses were in fact contaminated with mold.

Knowing which houses are contaminated and to what degree, as well as whether the residents are indeed experiencing mold-exposure-related health effects, will help health care workers diagnosis patients and will also help community leaders make evidence-based financial allocations for repairing houses or relocating residents living in band-owned houses. Our priority was to study human exposure to mold and the health effects of that exposure. The

work was done in three parts. First, the community's entire housing stock was inspected to determine whether houses were contaminated and to measure the amount of mold. Secondly, a questionnaire was administered to community residents to determine whether they were suffering any mold-related health effects. Thirdly, an attempt was made to create a predictor model to determine the amount of mold to which a person would have to be exposed before he or she would start to experience related health effects.

The main goal of the study was to attempt to answer the following three questions:

- How many houses contain visible mold and how much of it do they contain?

- How many residents are experiencing mold-related health effects, and what are their symptoms?

- What level of mold exposure is required to cause people to experience related health symptoms?

The expected benefit of this study in this Aboriginal community was that the houses that contained mold would be identified and the extent of contamination would be determined. This would help in prioritizing house repairs. Also, the inspection results could help health care professionals determine whether a patient's health symptoms were related to mold exposure at home. It was expected that the study would reveal that a number of houses were contaminated with mold and that some individuals were experiencing health symptoms because of exposure to mold in their homes.

Many residents of this community had indicated to the chief and council that they believed that their house was contaminated with mold and that they were concerned for their health.

Community health officials had attempted to respond to individual concerns by inspecting some houses for mold on request. However, they are not qualified inspectors and were not always confident about their findings. This inexperience and uncertainty did little to reassure homeowners, and there was considerable anxiety regarding mold contamination in the community, as many people were not sure whether their houses contained mold and whether their health was at risk. Since the community can be accessed only by air and since there are no qualified inspectors in the community, little had been done to address the issue. The Neskantaga Band had approached me to help address the situation.

3. Related Work and Background

Molds are ubiquitous in the environment; in fact, we are exposed to them every day, indoors and outdoors. Molds are classified as Myceteae (Fungi kingdom) (Alexopoulos et al., 1979). The Fungi kingdom includes molds, rusts, smuts and mushrooms (Alexopoulos et al., 1979). "Mold" is the common term for fungi that grow as a mat of entangled microscopic filaments known as hyphae (Hardin, 2002). Because molds do not have chlorophyll for energy production, they are considered parasitic organisms, dependent on external sources of food (Alexopoulos et al., 1979). They reproduce via sexual and asexual production of spores (Alexopoulos et al., 1979). Their classification is complex and is the subject of much debate among mycologists, owing mostly to the fact that there is a large variety of these organisms. However, the most common mold classification system divides fungi into three major divisions (Alexopoulos et al., 1979):

- Gymnomycota: cellular molds
- Mastigomycota, the "lower fungi": water molds
- Amastigomycota, the "true fungi": yeasts, molds, mildews, rusts, smuts, puffballs and mushrooms

Over 100,000 mold species have been identified to date, and some scientists have estimated that a quarter of the world's biomass is composed of mold (Miller, 1992). It is nearly impossible to isolate oneself from molds, owing to their numbers and omnipresence, unless stringent air filtration is practised and isolation and sanitary measures are taken, as is done in hospital units that care for patients who receive organ transplants (Hardin, 2002). Fungi are natural organisms in the environment and grow freely outdoors. Mold spores can get indoors through open doors and windows, ventilation and air-conditioning systems or any other building opening. According to the United States Centers for Disease Control and Prevention, people and pets can also serve as vehicles for entry of mold spores into a building, as the spores can be carried on clothes, fur and shoes, among other things (CDC, 2002).

Mold growth in the indoor environment

When mold is allowed to move freely from the outdoor environment to the indoor environment on a continual basis (through an open window or door), the indoor mold will reflect the mold normally found outdoors (Tobin, 1987). However, given the right environmental conditions, mold may start growing and proliferating in the indoor environment (Hardin, 2002). Several studies have been conducted to determine the environmental conditions required for mold to grow indoors.

The general consensus is that, in order to develop and propagate indoors, mold requires oxygen, nutrients, water and a temperature between 4° and 37°C (NAHB, 2002a).

Specifically, mold requires an adequate substrate on which to cling and off which to feed, such as most cellulose-based materials, including, but not limited to, drywall, paper and paper products, ceiling tiles and wood. As water is required for mold growth, buildings with high humidity levels (80% or more for sustained periods (IEA, 1990)) or subject to chronic water damage are particularly at risk for mold contamination (Husman, 1996; Macher, 1999).

Inorganic materials such as dust, paints, wallpaper, insulation materials, carpet, fabric and upholstery can also support fungal growth (APA, 2001).

Health Effects of Indoor Mold

The air we breathe is necessary for life, but, if it is contaminated, it can also be a vehicle for diseases that can make us ill. Most molds found in the environment are not pathogenic to healthy individuals, but some may cause superficial infections involving the human skin (feet, groin, dry body skin), while others cause the more common allergic responses, such as allergic asthma or allergic rhinitis (also known as hay fever) (Hardin, 2002). Individuals who are immune-compromised are at significant risk for opportunistic fungal infections and may experience the most severe forms of these infections, such as hypersensitivity pneumonitis (Hardin, 2002).

It is estimated that 10% of the population has allergic antibodies to fungal antigens and that only half of these individuals can be expected to show clinical illness (Hardin, 2002).

Generally speaking, possible health effects associated with mold fall into three different categories:

1. allergic: sensitization and immune responses such as allergic rhinitis, asthma or hypersensitivity pneumonitis (Rom, 1983);
2. infectious: growth of the fungus in the body (Walker, 1998); and
3. toxic: disruption of cellular function and interaction with DNA (Ciegler et al., 1980).

Allergy

At least 70 allergens have been well characterized from spores or other parts of fungi (Kurup et al., 2002; Grony et al., 2002). Several studies (Brunekreef et al., 1989; Lieberman, 2003; Rea, 2003) have found a consistent association between mold in the indoor environment and respiratory problems such as persistent cough, chest illness, wheezing, shortness of breath and asthma. Other non-respiratory or systemic health effects reported by these studies include muscle pain, lethargy, sinusitis, headaches, gastrointestinal problems, insomnia, nausea, skin rashes, memory loss, rhinitis and anxiety/ depression/ irritability. These systemic effects have

not been documented as well, and their importance has not always been clearly established. The causes of these effects also need to be further determined; therefore, this is an area in which further studies are necessary (MIEH, 1999; Rylander, 1999; King et al., 2002).

Toxicity

Fungi produce a wide variety of toxic chemicals known as mycotoxins. Some are very potent carcinogens, while others are immunosuppressants. Some of the most common mycotoxins are aflatoxins, hepatotoxins and ochratoxins (Curtis et al., 2004). Some mycotoxins are reported to inhibit protein synthesis and to cause hemorrhage and vomiting (Curtis et al., 2004). The smell of molds comes primarily from volatile organic compounds (Curtis et al., 2004).

The fact that exposure to mycotoxins may adversely impact the health of individuals has been known since the early 20th century. However, the pathway by which mycotoxins cause injury (inhalation) is still being debated today (Curtis et al., 2004).

Infection

Some fungi are known to infect immunocompetent individuals, usually resulting in minor skin infections. Common fungi that may infect immunocompetent individuals are *Candida*, *Histoplasma*, *Trichophyton* and *Malasezia* (Johnson et al., 1989). However, the most serious infections, caused by fungi such as *Candida*, *Aspergillus* and *Pneumocystis*, are usually seen in severely immunocompromised individuals (Nicod et al., 2001; Garber, 2001). Patients who undergo lung transplants, allogenic bone marrow transplants or heart transplants, or who suffer from acute leukemia are the most likely to suffer from invasive fungal infections, such as aspergillosis (Denning, 1998; Kontoyannis et al., 2002). Life-threatening fungal infections have been on the rise in recent years and the condition is growing rapidly (Husain et al., 2003).

Common Confounders

Exposure to several other substances commonly present in indoor environments can confound the association between mold exposure and patients' health symptoms. Such exposure needs to be ruled out in any situation where an individual presents with what appear to be mold exposure health effects. The common non-fungal environmental factors include poor ventilation, carbon monoxide from faulty heat sources, leaking natural gas, pesticides, wood-burning smoke, second-hand tobacco smoke, petrochemicals such as cleaners or solvents, and formaldehyde from off-gassing carpets (Curtis et al., 2004). Common household animal-related environmental factors include allergens from feathers, fur, saliva and excrement of cockroaches, dust mites, cats, dogs, mice, rats and caged birds (Higgins et al., 2000; Fireman, 2002).

Mold Prevention and Remediation

When visible fungi start growing indoors, every effort should be made to remove them and remediate the situation that caused mold to grow in the first place. Although surface mold on wood or other building materials does not cause structural damage, mold on wood and building materials indicates high moisture and adequate mold growth conditions that may promote the growth of wood decay fungi (NAHB, 2002b). That is one reason why it is important to limit mold growth in the indoor environment (APA, 2001).

Also, given the above-mentioned evidence that mold exposure is associated with certain specific health problems, removal of mold found in the indoor environment could lead to improved health in those experiencing mold-related health effects (Portnoy, 2005). Professional cleaning help may be required for buildings that are heavily contaminated with mold, but a few small areas of mold growth in a house can be taken care of by homeowners. Successful mold removal and remediation in a house contaminated with mold requires finding the mold and eliminating its source (Portnoy, 2005).

Moisture control is considered the best tool when trying to prevent mold growth. Moisture is created by daily activities such as cooking and showering, and this in turn creates microclimates in which humidity is higher, promoting mold growth. Therefore, mold may grow in these areas in an otherwise mold-free house. Mold growing in these areas can encourage mold proliferation in other areas of the house as well (Portnoy, 2005). Listed below are a few actions that can be taken to eliminate indoor moisture problems (Portnoy, 2005):

1. Maintain relative humidity levels at less than 50% indoors.
2. Seal all leaks to prevent water accumulation.
3. Use exhaust fans in the bathroom and kitchen to increase ventilation.

4. Vent dryer exhaust to the outside.
5. Limit the number of household plants that need watering.
6. Use air conditioners in the summer, when humidity levels are high.
7. Heat all rooms in the winter.
8. Use dehumidifiers in damp areas, such as basements.
9. If a house is prone to flooding, install sump pumps in the basement.

Carpets, wallpaper, paneling, and heating and air-conditioning systems are known to harbour fungal spores (Barr, 1999). Frequent vacuuming of carpets can help reduce mold spore levels, but replacing the carpet with hardwood tiles is a much more effective way to reduce spore levels, especially if the carpet is heavily contaminated with mold (Portnoy, 2005). For the treatment of nonporous surfaces such as wallpaper or paneling, a diluted bleach and detergent solution can successfully kill growing fungi and prevent their reappearance (Portnoy, 2005). Also, washable surfaces can be treated with commercial fungicidal compounds or antibacterial products specifically for mold and fungi. Some porous materials may have to be removed completely and replaced during remediation of severe mold contamination. Cleaning contaminated air ducts and filters may reduce mold exposure as well (Garrison et al., 1993). When an indoor environment is severely contaminated with mold, professional help may be required in order to remove the mold and remediate the problem. If severe mold contamination is present in a house or other building, mold removal experts will often use asbestos remediation techniques for removing the mold.

Mold is a part of everyday life and we can expect to be exposed to it everyday, whether indoors or outdoors. Mold may become problematic if it is allowed to grow unchecked indoors, to the point where it becomes visible. Small amounts of mold in a building are unlikely to produce severe health effects in those exposed, but large amounts of mold growth in the indoor environment may cause certain health effects in some individuals.

It is best to prevent mold from growing in the first place. If conditions that promote mold growth, such as high humidity or water leaks, do not exist in the indoor environment, mold is unlikely to grow. Therefore, prevention is the best way to avoid mold-related problems. Regular inspections should be conducted throughout the house or building to look for the presence of mold or any factors that could lead to mold growth.

Many studies have successfully linked indoor mold to a variety of respiratory and non-respiratory health effects in exposed individuals, although the mechanism by which mold causes the health effects is still not fully understood. Also, individuals vary in their susceptibility to development of mold-related health effects when exposed.

4. Methods

This project was executed in three stages. During the first stage, a team of qualified housing inspectors and trained community participants inspected every home in the community for mold. The location of any mold found was documented, and the amount of mold was measured. All areas of the houses were inspected, from attic to basement.

During the second stage, a team of experienced surveyors and trained community-based participants administered an approved health questionnaire to community members, in order to report possible mold-exposure-related health effects experienced by these community residents.

The third stage involved bringing the first and second stages together in an attempt to develop a predictor model. The purpose of developing such a model was to determine how much mold a person needs to be exposed to before he or she will start to experience mold-related symptoms. The model was to be created using an SPSS statistical database.

All inspections and surveys were done in the community of Neskantaga. Facilities used in the community were the community's band office for office supplies and services, the community hall for meetings, and the community radio station to inform the community of project-related issues.

During the study, a project leader oversaw the entire process. A certified housing inspector was responsible for training 5 community members in how to conduct the housing inspections. The certified inspector oversaw the housing-inspection stage.

There was also an experienced survey administrator who was responsible for training 5 community members in how to conduct the survey questionnaire in the community.

Both the housing inspector and the health questionnaire administrator prepared mock activities to ensure that the trainees were adequately trained and prepared to do house inspection and questionnaire administration on their own and without direct supervision.

5. Data analysis

Table 3: Definition of dependent and independent variables

5.1 Detailed file by individual (*N=193*)

Dependent Variables (names of variables in italics)	
Sum of physical symptoms (*Sum_physical_sym*)	Sum of the following variables: *Sinusitis, Bronchitis, Pneumonia, Asthma, Cough, Phlegm, Wheezing, Emphysema* and *Rhinitis*, regardless of frequency
Sum of sensitivity symptoms (*Sum_sensitivity_sym*)	Sum of the following variables: *WateryEyes, DrySkin, DryEyes, DryNose* and *DryThroat*, regardless of frequency
Sum of subjective symptoms (*Sum_subjective_sym*)	Sum of following variables: *Lethargy, Headaches, Malaise* and *SkinRash*, regardless of frequency
Sum of all symptoms (*Sum_overall_sym*)	Sum of the preceding subtotals: *Sum_physical_sym*, *Sum_sensitivity_sym* and *Sum_subjective_sym*
Physical symptoms	Physical symptom reported or not (non-transformed variables, coded *1*=yes and *0*=no: *Sinusitis*, *Bronchitis*, *Pneumonia*, *Asthma*, *Cough*, *Phlegm*, *Wheezing*, *Emphysema* and *Rhinitis*)
Sensitivity symptoms	Sensitivity symptom reported or not (non-transformed variables, coded *1*=yes and *0*=no: *WateryEyes, DrySkin, DryEyes, DryNose* and *DryThroat*)
Subjective symptoms	Subjective symptom reported or not (non-transformed variables, coded *1*=yes and *0*=no: *Lethargy, Headaches, Malaise* and *SkinRash*)
Independent Variables (names of variables in italics)	
Mold	Number of square feet inside the house where mold is visible (*Mold* – non-transformed variable)
Gender of the respondent	*Gender* = non-transformed variable (*0*=male, *1*=female)
Presence of a dog, cat, bird or other animal (*Dog*, *Cat*, *Bird* and *Other*)	Created from the *PetType* variable, if there is at least one domestic animal in the home (*PetInHouse* equals zero)
Presence or not of pets (*Pets*)	Created by reversing the coding of *PetsInHouse* (*1*=yes, *0*=no)
Occupation outside the home (*OccupiedOutside*)	Created by dichotomizing the occupation type (*Occupation*) on the basis of whether the occupation frequently takes the person out of the home (the values *Baby, Disability, Elder, Homecare, Housekeeper, Self-employed, Toddler* and *Unemployed* produce a value of *0*, whereas all other occupations produce a value of *1*)

Smoker (*Smoke*)	Created by reversing the coding of *Smoke* (*1*=yes, *0*=no)
Level of smoking (*SmokeFrequency*)	Created by modifying the coding of *SmokeFrequency* so that non-smokers are assigned a value of *0* instead of *5*, while the *1* to *4* categories continue to represent an increasing level of smoking
Number of people living in the home (*PersonsLiving*)	Created by counting the number of respondents per home

Presence of a smoker in the home (*Any_smoker*)	Created by identifying the presence of at least one smoker in the home, on the basis of the *Smoke* variable (*0*=no smoker, *1*=one or more smokers)
At least one person smokes inside the home (*Any_smoker_inside*)	Created by identifying the presence of at least one person smoking inside the home, on the basis of the *SmokeInHouse* variable (*0*=no smoker inside the home, *1*= one or more smokers inside the home)
Number of people who smoke inside the home (*N_smokers_inside*)	Created by counting the number of people smoking inside the home, on the basis of the *SmokeInHouse* variable (*0*= no smoker inside the home, *1* to *3*=number of people smoking inside the home)

5.2 Detailed file by home/household (*N=59*)

Dependent Variables (names of variables in italics)	
Average number of physical symptoms (*AvgN_physical_sym*)	Average number of physical symptoms in the home (variable *Sum_physical_sym* described above)
Average number of sensitivity symptoms (*AvgN_sensitivity_sym*)	Average number of sensitivity symptoms in the home (variable *Sum_sensitivity_sym* described above)
Average number of subjective symptoms (*AvgN_subjective_sym*)	Average number of subjective symptoms in the home (variable *Sum_subjective_sym* described above)
Average number of symptoms (all types) (*AvgN_overall_sym*)	Average of the total number of symptoms (variable *Sum_overall_sym* described above)
% afflicted with each physical symptom	% of people in the home afflicted with each physical symptom (*Sinusitis*, *Bronchitis*, *Pneumonia*, *Asthma*, *Cough*, *Phlegm*, *Wheezing*, *Emphysema* and *Rhinitis*)
% afflicted with each sensitivity symptom	% of people in the home afflicted with each sensitivity symptom (*WateryEyes, DrySkin, DryEyes, DryNose* and *DryThroat*)
% afflicted with each subjective symptom	% of people in the home afflicted with each subjective symptom (*Lethargy, Headaches, Malaise* and *SkinRash*)
Independent Variables (names of variables in italics)	

Mold	Number of square feet in the house where mold is visible (*Mold* – non-transformed variable)
Respondent is a regular smoker (*Smoke*)	Untransformed variable (*1=yes, 0=no*)
Number of people living in the home (*PersonsLiving*)	Created by counting the number of respondents per home
% of young children, children/adolescents, and seniors (*PctBabies, PctChildren, PctElder*)	% of people living in the home who are young children (5 years and under), children or adolescents (17 years and under) or seniors (60 years and over)
Presence or not of pets (*Pets*)	Created by reversing the *PetsInHouse* coding (*1*=yes, *0*=no)
Presence or not of a dog (*Dog*)	Created from the *PetType* variable
Presence of a smoker in the home (*Any_smoker*)	Created by identifying the presence of at least one smoker in the home, according to the *Smoke* variable (*0*=no smoker, *1*=one or more smokers), whether or not this individual smokes inside the home
At least one person smokes inside the home (*Any_smoker_inside*)	Created by identifying the presence of at least one person smoking inside the home, according to the *SmokeInHouse* variable (*0*=no one smoking inside the home, *1*=one or more persons smoking inside the home)
Number of people smoking inside the home (*N_smokers_inside*)	Created by counting the number of people smoking inside the home, according to the *SmokeInHouse* variable (*0*=no one smoking inside the home, *1* to *3*=number of people smoking inside the home)
% of residents of the home who spend a large part of the day outside the home (*OccupiedOutside*)	Created by dichotomizing the occupation type (*Occupation*) using the above-described approach, then calculating the percentage of residents of the home who spend a lot of time outside the home

Although identified separately, symptoms were analysed only on the basis of their presence/absence. Therefore, the frequency of these symptoms is not known, as the frequency of some symptoms was not evaluated.

Note that several of these analyses were carried out on the basis of aggregate data by home, in addition to being carried out on the basis of detailed data by respondent. The results of these analyses are reported in the appendix. This approach was chosen because it is not unusual for individuals living in the same home to report different incidences of symptoms, even though they are exposed to the same environment as other members of the household. Therefore, the

observations are not strictly independent of one another, as common quantitative statistical methods require. A comparison of the analyses by household with those of individuals reveals that similar results were obtained.

6. Results

6.1 Relationship between mold and symptoms

Because the relationship between the level of mold and the symptoms reported by respondents is at the core of the research objectives, it is appropriate to first present the Pearson correlations between each symptom or group of symptoms and the mold level.

Using the detailed data by respondent (n=193), we obtain the results presented in Table 1.

Table 4: Pearson correlations between the dependent variables and the quantity of mold (detailed data for each respondent)

Dependent variable	Correlation
Sum of physical symptoms	-.01
Sum of sensitivity symptoms	.07
Sum of subjective symptoms	.12
Sum of all symptoms	.06
Sinusitis	-.10
Bronchitis	-.02
Pneumonia	-.07
Asthma	-.07
Cough	.13
Phlegm	.05
Wheezing	.00

Dependent variable	Correlation
Rhinitis	-.02
Watery eyes	-.07
Dry skin	.20 *
Dry eyes	.20 *
Dry nose	-.10
Dry throat	-.02
Lethargy	.16 *
Headaches	.07
Malaise	.12
Skin rash	-.01

* Significant correlations at the 5% threshold are indicated with an asterisk.

There is no universally accepted threshold for assessing Pearson correlations; they vary greatly, depending on the field of study. However, there is general agreement in social sciences and epidemiology that interesting correlations start at *0.30*. The Pearson coefficient is definitely affected marginally by the fact that individual symptoms are measured on a binary scale (presence/absence), but this does not invalidate the results obtained. In this analysis, there seems to be no strong correlation between the quantity of mold in a house and the reported symptoms. There is a correlation between mold and dry skin, dry eyes and lethargy, but this correlation remains weak.

6.2 Symptom prediction

Multiple regressions were used to construct predictive models of the number of symptoms reported. To predict each of the symptoms individually, logistic regression was used. The reason for this is that, when each symptom is considered in isolation, these dependent variables contain only data related to presence/absence, and only logistic regression can be used under these conditions. (See the methodology appendix for further details.) The predictive variables used are listed in Table 3.

Some predictive variables had to be rejected. The source of heat in the dwelling might have revealed interesting elements, but it was constant in the sample studied: all homes were heated with wood. The type of pet is another indicator that seems relevant, but dogs were the only pets present in the sample. Moreover, the number of pets might have been interesting, but was not recorded.

Table 5: Predictive variables

Predictive variables used
Gender of the respondent
Age of the respondent
Number of square feet of mold
Number of people living in the home
% of young children in the home
% of children and adolescents in the home
% of seniors in the home
Presence or absence of pets in the home
Presence or absence of a dog in the home
Respondent is a smoker
Respondent's level of smoking
Presence of one or more smokers in the home (smokes or does not smoke inside the home)
Number of people who smoke inside the home
% of residents who spend a large part of the day outside the home

6.2.1 Continuous dependent variables

Table 5 summarizes the multiple regression models identified. The R^2 column contains the R squared—that is, the proportion of variance of the dependent variable that can be explained by the regression model (this value varies from *0 to 1*; the higher the number, the better the

model's explanatory value). Columns in the predictor block contain coefficients (non-standardized) that must be applied to each independent variable to explain the dependent variable. Because a multiple regression model was used, each line of the table corresponds to a series of coefficients used to predict one of the dependent variables. Only predictors with a 5% level of significance are presented in the table; the other predictors (see Table 3) were not retained in these models.

Table 6: Multiple regression models for continuous dependent variables

Dependent variable	R^2 (proportion of variance explained)	Predictors			
		Constant	Age	At least one smoker living in the home	N of people in the home
Sum of physical symptoms	.10	1.84	---	-1.15 (-0.24)[1]	0.32 (0.26)
Sum of sensitivity symptoms	.18	0.95	0.018 (0.21)	-1.05 (-0.26)	0.28 (0.27)
Sum of subjective symptoms	.26	0.21	0.022 (0.30)	-0.83 (-0.25)	0.30 (0.35)
Sum of all symptoms	.23	2.48	0.049 (0.22)	-2.89 (-0.27)	0.94 (0.35)

Note 1: The coefficient is on the first line, and the correlation (semi-partial) between the predictor and the dependent variable is on the second line.

All the prediction models for the quantitative dependent variable produce low proportions of explained variation (10 to 26%). There is no universally accepted standard, but, in most social science and epidemiological studies, a model begins to be of interest when it is responsible for explaining 30% (preferably more) of the variance in the dependent variable.

Note that the relative magnitude of coefficients depends greatly on the variability of the predictor and cannot be used to assess the significance of one predictor in relation to others. Thus, since *Age* varies a great deal more than *At Least One Smoker In The Home*, its coefficient is naturally lower, all other things being equal. The relative "strength" of a predictor is assessed here by using its semi-partial correlation with the dependent variable (clean correlation of the relationships between predictors).

Note that *Age* is a key element in three of the four models and that it has a positive coefficient. The older the person, the more he or she reports symptoms. However, for some reason that I cannot explain, the number of symptoms reported is inversely related to the presence of smokers in the home: homes with non-smokers report more symptoms than homes with smokers. The number of people living in a home is another predictor that was retained. Semi-partial correlations indicate that predictors all have about the same significance in each model.

6.2.2 Binary dependent variables

Table 7 presents the chi-squares for each of the analyses, as well as symptom attack, sensitivity and specificity rates.

In all cases, a significant chi-square indicates that the model, with its predictors, is superior to an empty model—that is, a model with no predictor (if it is not the intercept). All symptoms can therefore be predicted with a success rate (sensitivity) that is clearly variable. The only exception is *Rhinitis*, for which we were unable to identify any predictor.

Table 7: Chi-square results, sensitivity/specificity and correct classification rates

Dependent variable[1]	% sensitivity/ % specificity	Correct classification rate (%)	Chi-square	Prob.
Sinusitis	53/76	58	15.2	0.001
Bronchitis	77/57	75	10.2	0.006
Pneumonia	67/55	65	6.2	0.013
Asthma	93/21	81	9.5	0.009
Cough	59/52	55	8.5	0.015
Phlegm	61/60	61	14.7	0.002
Wheezing	74/47	66	19.7	0.000
Watery eyes	81/65	76	49.6	0.000
Dry skin	73/46	63	7.7	0.005
Dry eyes	68/52	63	13.1	0.001
Dry nose	65/69	66	30.9	0.000
Dry throat	63/58	61	8.9	0.012
Lethargy	73/72	73	54.4	0.000
Headaches	72/71	72	43.1	0.000
Malaise	70/73	71	36.5	0.000
Skin rash	68/54	65	10.9	0.004

Note 1: The *Emphysema* and *Rhinitis* variables were not analysed. (No subjects reported emphysema, and, for rhinitis, we were unable to enter any predictor in the model.)

In the context of this study, it is certainly more interesting to focus on sensitivity rates—that is, on the ability to correctly identify subjects suffering from a symptom (true positives). The specificity rate (correctly identifying subjects not suffering from the symptom, or "true negatives") is more useful in a clinical context.

It is utopian to think that we can have both sensitivity and specificity equal to 100%, even in ideal conditions and with perfect identification of all the predictors involved in the occurrence of a symptom. Also, there are no precise criteria to indicate where a sensitivity rate or sufficient specificity begins, provided, of course, that it exceeds a value of 50%, which represents pure chance.

Note that there were sensitivity rates of over 70% for the following variables: *Bronchitis*, *Asthma*, *Wheezing*, *WateryEyes*, *DrySkin*, *Lethargy*, *Headaches* and *Malaise*. That is very promising. An examination of the dependent variables involved indicates that they are not necessarily the ones with the easiest symptoms to identify (think of *Lethargy* and *Malaise*, the definitions of these words no doubt vary greatly from one individual to the next).

Table 8 presents the logistic regression coefficients and corresponding relative risks. These coefficients are similar to B coefficients in standard linear regression and may be useful for establishing a prediction equation and thus generating predicted values. (For example, study another sample of Aboriginal people and use an examination of the dwelling and lifestyle of each of them as a basis for predicting the occurrence of the health symptoms studied here.) The "A" intercept in the equation describes a model that does not include any of the predictive variables and, although reported here, is not very useful in our context.

It is immediately obvious that the *Any_Smoke*, *Smoker* and *PersonsLiving* variables show up most often with significant risk ratios in our models. It is interesting to note that, for each increase of 1 in the *Smoker* variable, the risk is multiplied by more than 2.5 for the *Sinusitis*, *Phlegm*, *DryNose* and *Malaise* variables, when the influence of all other variables involved in the model is kept constant. Moreover, note that increasing age resulted in a higher prevalence of symptoms in a marginal but regular way: an increase of 10 years multiplied by 1.2 to 1.5 the risk of suffering from certain symptoms. Also, note that the *Pets* variable multiplies the risk of *Asthma* by more than 3.35. Finally, being a woman doubles the risk of *Headaches*.

In summary, although the predictive capacity of logistic regression models is sometimes limited, it is possible to explain part of each of the symptoms (except *Rhinitis*) using the selected predictors. A finer analysis, conducted symptom by symptom, seems more promising.

Table 8: Logistic regression models for binary dependent variables

Dependent variable[1]	Statistic	Constant	At least one smoker living in the home	N of people in the home	Pets	Smoker	Mold	Age	Gender
Sinusitis	Coefficient	-0.37	-1.70	---	---	0.97	---	---	---
	Relative risk[2]	0.69	0.18	---	---	2.63	---	---	---
Bronchitis	Coefficient	-2.65	-1.04	0.31	---	---	---	---	---
	Relative risk	0.07	0.36	1.37	---	---	---	---	---
Pneumonia	Coefficient	-3.55	---	0.33	---	---	---	---	---
	Relative risk	0.03	---	1.40	---	---	---	---	---
Asthma	Coefficient	-1.66	---	---	1.21	---	---	---	---
	Relative risk	0.10	---	---	3.35	---	---	---	---
Cough	Coefficient	0.07	-0.78	0.24	---	---	---	---	---
	Relative risk	1.07	0.46	1.27	---	---	---	---	---
Phlegm	Coefficient	-1.33	-1.09	0.26	---	0.94	---	---	---
	Relative risk	0.27	0.34	1.30	---	2.56	---	---	---
Wheezing	Coefficient	-1.63	-1.17	0.37	---	---	---	---	---
	Relative risk	0.20	0.31	1.45	---	---	---	---	---
Watery eyes	Coefficient	-1.86	-1.54	0.27	---	---	---	0.04	---
	Relative risk	0.16	0.21	1.32	---	---	---	1.04	---
Dry skin	Coefficient	-1.28	-0.93	0.30	---	---	0.02	---	---
	Relative risk	0.28	0.39	1.35	---	---	1.02	---	---
Dry eyes	Coefficient	-0.46	-0.88	---	---	---	0.03	---	---

	Relative risk	0.63	0.42	---	---	---	1.03	---	---
Dry nose	Coefficient	-0.82	-1.60	0.29	---	0.90	---	---	---
	Relative risk	0.44	0.21	1.33	---	2.46	---	---	---
Dry throat	Coefficient	-1.76	---	0.16	---	---	---	0.02	---
	Relative risk	0.17	---	1.17	---	---	---	1.02	---
Lethargy	Coefficient	-2.77	-1.50	0.54	---	---	0.02	0.04	---
	Relative risk	0.06	0.22	1.72	---	---	1.02	1.04	---
Headaches	Coefficient	-3.10		0.35	---	---	---	0.05	0.70
	Relative risk	0.05		1.42	---	---	---	1.05	2.02
Malaise	Coefficient	-2.50	-1.95	0.48	---	0.90	---	0.02	---
	Relative risk	0.09	0.14	1.61	---	2.45	---	1.02	---
Skin rash	Coefficient	-1.39	-1.22	0.19	---	---	---	---	---
	Relative risk	0.25	0.30	1.21	---	---	---	---	---

Note 1: The *Emphysema* and *Rhinitis* variables were not analysed. (No subjects reported emphysema, and, for rhinitis, we were unable to enter any predictor in the model.)
Note 2: Relative risk or "odds ratio."

7. Conclusion and recommendations

The study was successful in a variety of ways, as we were able to inspect all the houses in the community and determine not only the number of houses that contained visible mold, but also the amount of mold found. This information is invaluable to the community leaders: they now have a detailed map of the mold situation in the community and can undertake decontamination, giving priority to the houses in greater need of remediation. We were also able to determine how many residents were experiencing mold-related health effects. This information, coupled with the recorded measurements of mold in the houses, can help health care providers formulate a better diagnosis for those individuals who seek medical attention.

However, one of the main goals of the study—that of determining the level of mold to which a person needs to be exposed to experience related health symptoms—was not achieved conclusively. The *Mold* variable was absent in all the models (both multiple and logistic regression); therefore, in this study, it was not a reliable symptom predictor. However, we must not jump to conclusions and state that it has no effect on symptoms: the analysis should probably be taken further and the subject investigated from another angle.

In this regard, we would like to make a few recommendations regarding research methodology, with a view to increasing the chances of obtaining robust and predictive models. When measuring contamination, if the area of dwellings varies substantially among dwellings in a sample, ratios that take into account the area, rather than a measurement of the contaminated surface, should be favoured. Also, data should be collated on the frequency of all symptoms and their progression should be clarified, so that these variables can be used in the models. An indicator of the percentage of time spent by each person outside the dwelling during a typical week should be recorded, since the elderly, stay-at-home mothers, babies, people with physical disabilities and self-employed workers are likely to spend much more time in the dwelling than people working outside the home, such as students, children and adolescents. An indicator of the ambient humidity level should be recorded, since indoor moisture is related to mold growth. Also, frequency measures should be refined when one category contains the vast majority of respondents.

Although the main objective of creating a predictor model was not attained, I hope that the investigation will result in increased awareness of the hazards of living in a house contaminated with mold. In addition, I hope that some capacity building has resulted at the community level, since individuals were trained in identifying mold and were educated regarding cleanup measures to take when the amount of mold permits remediation by the homeowner. Also, because the houses contaminated with mold have been identified, the community leaders have an opportunity to tackle the mold problem by prioritizing house remediation on the basis of mold content. Removal of the exposure source might reduce the mold morbidity burden.

8. References

1. Alexopoulos CJ, Mims CW. Introductory Mycology, 3rd ed. John Wiley & Sons, New York. 1979

2. APA – The Engineered Wood Association.. Mold and Mildew. 2001

3. Barr M. Indoor fungal exposure: does it matter, and what can be done about it? Clin Exp Allergy. 1999; 29: 1442-1444

4. Bobbit RC Jr. Characterization of a population presenting with suspected mold-related health effects. Ann Allergy Asthma Immunol. Jan 2005; 94, 1

5. Brunekreef B, Dockery DW, Speizer FE, Ware JG, Spengler J, Ferris BG. Home dampness and respiratory morbidity in children. Am Rev Respir Dis. 1989; 140: 1353-67

6. CDC National Center for Environmental Health. 2002 Questions and Answers on *Stachybotrys chartarum* and other molds. Accessed 2005 October 7

7. Ciegler A, Bennet JW. Mycotoxins and Mycotoxicoses. BioScience. 1980; 30: 512-515

8. Curtis L, Lieberman A, Stark M, Rea W, Vetter M. Adverse Health Effects of Indoor Molds. Journal of Nutritional & Environmental Medicine. 2004; 14(3): 261-274

9. Denning D. Report on a European Science Foundation Workshop on Invasive Aspergillosis, 21-22 October 1998, University of Manchester, Manchester

10 Fireman P. Virus-provoked rhinitis in patients who have allergies. Allergy Asthma Proc. 2002; 23(2): 99-102

11. Garber G. An overview of fungal infections. Drugs 2001; 61(Supp 1): 1-12

12. Garrison R, Robertson L, Koehn R, Wynn S. Effect of heating-ventilation-air conditioning system sanitation on airborne fungal populations in residential environments. Ann Allergy. 1993; 71: 548-556

13. Grony RL, Reponen T, Willeke K. Fungal fragments as indoor air contaminants. Applied Environmental Microbiology. 2002; 68: 3522-31

14. Hardin BD.; Adverse human health effects associated with molds in the indoor environment. American College of Occupational and Environmental Medicine. 2002

15. Higgins BG, Francis HC, Yates G. Environmental exposure to air pollution and allergens and peak flow changes. European Respiratory Journal. 2000; 16(1): 61-6

16. Husain S, Alexander BD, Munoz P. Opportunistic mycelial fungal infections in organ transplant recipients: emerging importance of non-Aspergillus fungi. Clin Infect Dis. 2003; 37(2): 221-9

17. Husman T. Health effects of indoor air microorganisms. Scand J Work Environ Health. 1996; 22:5-13

18. International Energy Agency. Annex 14: Condensation and Energy, Vol 2: Guidelines and Practice. 1990

19. Johnson P, Sarosi G. Community-acquired fungal pneumonias. Semin Resp Infect. 1989; 4(1): 56-63

20. King, N, Auger, P. Indoor air quality, fungi, and health: How do we stand? Canadian Family Physician – Le Médecin de famille canadien, Vol 48, February 2002

21. Kontoyannis D, Bodey G. Invasive aspergillosis in 2002: an update. Eur J Clin Microbiol Infect Dis. 2002; 21: 161-72

22. Kurup V, Shen HD, Vijay H. Immunobiology of fungal allergens. International Archives of Allergy Immunology. 2002; 129: 181-8

23. Lieberman A. Explosion of mold cases in homes, workplaces and occupational medicine practices. Presented at the 21st Annual Symposium on Man and his Environment in Health and Disease. Dallas, Texas. 19-22 June 2003

24. Macher J, editor. Bioaerosols: assessment and control. Cincinnati, Ohio: American Conference of Governmental Industrial Hygienists. 1999

25. McMaster Institute of Environmental Health (CAN). Expert Panel on Fungal Contamination Indoors. Hamilton, Ontario: McMaster Institute of Environment and Health. 1999

26. Miller, JD. Fungi as contaminants of indoor air. Atmos Environ. 1992; 26A(12): 2162-72

27. NAHB Research Center's ToolBase Services. 2002. Helping Your Buyers Understand Mold During the Building Process

28. National Aboriginal Health Organization. Regional Longitudinal Health Survey, 2002-2003

29. National Association of Home Builders Research Center's Toolbase Services. Mold in Residential Buildings. 2002

30. Nicod L, Pache J, Howarth N. Fungal infections in transplant recipients. European Respiratory Journal. 2001; 17(1): 133-40

31. Portnoy MJ, Kwak K, Dowling P, VanOsdol T, Barnes C. Annals of Allergy, Asthma and Immunology. Palatine: Mar 2005. Vol 94, Issue 3: p 313

32. Rao, CY, Burge, HA, Chang JCS. Review of Quantitative Standards and Guidelines for Fungi in Indoor Air. J Air Waste Manage Assoc. 1996; 46: 899-908

33. Rea WJ, Didriksen N, Simon TR, Pan Y, Fenyves EJ, Griffiths B. Effects of toxic exposure to molds and mycotoxins in building-related illnesses. Arch Environ Health. 2003; 58(7): 399-405

34. Report of the Auditor General of Canada, 2003

35. Robbins AC, Swenson JL, Nealley LM. Health Effects of Mycotoxins in Indoor Air: A Critical Review. Applied Occupational and Environmental Hygiene. 2000; Vol 15(10): 773-784

36. Rom WN. Environmental and Occupational Medicine. Little, Brown and Co. Boston. 1983

37. Rylander R, Etzel R. Introduction and summary: workshop on children's health and indoor mold exposure. Environmental Health Perspective. 1999; 107 (Supp 3): 465-8

38. Statistics Canada. 2008. *Neskantaga, Ontario* (table). *Aboriginal Population Profile.* 2006 Census. Statistics Canada Catalogue no. 92-594-XWE. Ottawa. Released January 15, 2008.

39. Tobin RS. Signification de la presence de champignons dans l'air à l'intérieur des édifices. Rev Can Santé Publique. 1987; 78(Supp 2): S15-S32

40. Walker TS. Microbiology. WB Saunders Co. Philadelphia.1998

Appendix A

Health Symptoms Questionnaire

Questionnaire

Interviewer's Full Name: _____

INTERVIEWER INTRODUCTION

Hello, my name is _____. As part of a community wide mold investigation, I am conducting a short survey. The survey takes about 10 to 15 minutes and deals with questions relating to your general health, with particular attention to your respiratory health.

IMPORTANT: IMMEDIATELY FILL IN DATE AND TIME STARTED

DATE_____ 20 __ TIME STARTED_____ ( )A.M. ( )P.M.

1. What is your full name?

2. What is your Age?

3. What is you gender?

Respiratory Health Questions

Sinusitis

4 a. Have you ever had sinusitis in the last 12 months?

O Yes O No

4 b. If yes, have you been diagnosed with sinusitis?

O Yes O No

Bronchitis

5 a. Have you ever had bronchitis in the last 12 months?

O Yes O No

5 b. If yes, have you been diagnosed with chronic bronchitis?

O Yes O No

Pneumonia

6 a. Have you ever had pneumonia in the last 12 months?

O Yes O No

6 b. If yes, have you been diagnosed with pneumonia?

O Yes O No

Asthma

7 a. Have you ever had asthma?

O Yes O No

7 b. If so, have you had an attack of asthma at any time during the last 12 months?

O Yes O No

7 c. Are you currently taking any medicines (inhalers, aerosols or pills) for asthma?

O Yes O No

Cough

8 a. During the last 12 months, have you usually had a cough?

O Yes O No

8 b. If so, do you usually cough as much as 4 to 6 times a day, 4 or more days out of the week?

O Yes O No

8 c. Do you usually cough like this on most days for 3 consecutive months or mor during the year?

O Yes O No

8 d. For how many years have you had this cough ?_____

Phlegm (mucus from chest)

9 a. Do you usually bring up phlegm form your chest?

O Yes O No

9 b. Do you usually bring up phlegm like this as much as twice a day, 4 or more days out of the week?

O Yes O No

9 c. Do you usually bring up phlegm like this on most days for 3 consecutive months or more during the year?

O Yes O No

9 d. For how many years have you had trouble with phlegm? _____

Wheezing

10 a. During the last 12 months, have you usually had a wheeze or whistle from your chest?

O Yes O No

10 b. If so, do you usually wheeze or whistle from your chest as much as 4 to 6 times a day, 4 or more days out of the week?

O Yes O No

10 c. Do you usually wheeze or whistle from your chest like this on most days for 3 consecutive months or more during the year?

O Yes O No

10 d. For how many years have you had this wheeze or whistle problem? _____

Emphysema

11. Has a doctor ever told you that you had emphysema?

O Yes O No

Other Health Symptoms

Rhinitis (runny nose)

12 a. During the last 12 months, have you usually had a runny nose?

O Yes O No

12 b. If so, do you usually have a runny nose as much as 4 to 6 times a day, 4 or more days out of the week?

O Yes O No

12 c. Do you usually have a runny nose like this on most days for 3 consecutive months or more during the year?

O Yes O No

12 d. For how many years have you had this runny nose problem? _____

Watery Eyes

13 a. During the last 12 months, have you usually had watery eyes?

O Yes O No

13 b. If so, do you usually have watery eyes as much as 4 to 6 times a day, 4 or more days out of the week?

O Yes O No

13 c. Do you usually have watery eyes like this on most days for 3 consecutive months or more during the year?

O Yes O No

13 d. For how many years have you had this watery eyes problem? _____

14. In the past 12 months, have you experienced any of the following SYMPTOMS?

a) Dryness of the skin O Yes O No

every day _____
every week _____
every month _____
other _____

b) Dryness of the eyes O Yes O No

every day _____
every week _____
every month _____
other _____

c) Dryness of the nose O Yes O No

every day _____
every week _____
every month _____
other _____

d) **Dryness of the throat** ○ Yes ○ No

every day _____
every week _____
every month _____
other _____

e) **Lethargy** ○ Yes ○ No

every day _____
every week _____
every month _____
other _____

f) **Headaches** ○ Yes ○ No

every day _____
every week _____
every month _____
other _____

g) **Malaise** ○ Yes ○ No

every day _____
every week _____
every month _____
other _____

h) **Skin rash** ○ Yes ○ No

every day _____
every week _____
every month _____
other _____

15. What heating source do you use to heat your house?

Wood stove _____
Oil furnace _____
Electricity _____
Gas _____
Other _____

16. Do you have any pet living inside the house?

O Yes O No

If yes, what kinds and how many of each? _____

17. Tobacco Smoking

17 a. Do you currently smoke?

O Yes O No

17 b. If yes, what have been our smoking habits during the last 12 months?

O Less than half a pack per day
O More than half a pack per day
O One pack per day
O More than one pack per day

17 c. Do you smoke inside the house?

O Yes O No

18. What has been your occupation(s) during the last 12 months?

_____
_____
_____

Comments / Questions

Appendix B

Mold Measurement Form

Investigator: _____ **Date:** _____

Occupants: 1. 4.
2. 5.
3. 6.

N/G=No Growth S=Small(<10 sq. ft) M=Moderate(10-30 sq. ft) L=Large(>30 sq. ft)

Room	**Calculation in Sq/ft**	**Specific Location or Comments**	**N/G**	**S**	**M**	**L**
Crawl Space/ Ground conditions						
Full Basement						
Kitchen						
Bathroom						
Living Room						
Master Bedroom						
2nd Bedroom						
3rd Bedroom						
4th Bedroom						
Laundry/Utility						
Closet						
Porch						
Other						
Total						

General Cleanliness:

Humidity %: _____

Temperature (Celsius):_____

Air Exchange Unit: Yes_____ **No**_____

Comments:_____

_____

_____

Appendix C

Tables and complementary results

The following tables present the analytical results obtained by using aggregate data by household. Note that the trends observed from detailed data by respondent continue to be present when the data is analysed by household.

Table A1 – Pearson correlations between the dependent variables and the quantity of mold (aggregate data by household)

Dependent variable	Correlation
Average N of physical symptoms	.00
Average N of sensitivity symptoms	.05
Average N of subjective symptoms	.12
Average N of symptoms (all types)	.06
Sinusitis	-.10
Bronchitis	.00
Pneumonia	-.08
Asthma	-.09
Cough	.16
Phlegm	-.06
Wheezing	-.01

Dependent variable	Correlation
Rhinitis	.00
Watery eyes	-.12
Dry skin	.27 *
Dry eyes	.21
Dry nose	-.13
Dry throat	-.04
Lethargy	.18
Headaches	.09
Malaise	.08
Skin rash	-.02

* Significant correlations at the 5% threshold are indicated with an asterisk.

Table A2 – Multiple regression models for quantitative dependent variables (aggregate data by household)

Dependent variable	R^2 (explained proportion of variance)	Predictors		
		Constant	At least 1 smoker living in the home	N of people in the home
Average N of physical symptoms	.11	1.98	-0.99	0.5
Average N of sensitivity symptoms	.07	2.25	-0.76	---
Average N of subjective symptoms	.16	1.94	-0.89	---
Average N of symptoms (all types)	.09	6.76	-2.28	---

Appendix D

How to interpret the results of a logistic regression

Respective fields for linear and logistic regression

Logistic regression is used in situations where we want to predict the presence or absence of a characteristic or event on the basis of the values of a set of predictive variables. In this way, it is similar to a linear regression model, but it is clearly more appropriate when the dependent variable is dichotomous.

With logistic regression, it is desirable to code the dependent variable with the value 1 when the symptom is present and with the value 0 when the symptom is absent, and this is the case in this report. Reversing the coding forces us to reverse several of our interpretations, adding to the complexity of the analysis.

With a dichotomous dependent variable, we cannot apply a traditional linear regression model for the following three reasons:

Linear regression does not place any constraint on the predicted values. We therefore will obtain predicted values above 1 and below 0, as well as fractional values between 0 and 1. None of this makes sense, because the dependent variable measured confirms the presence (1) or absence (0) of the symptom.

1. One of the postulates in linear regression is the reliability of the variance of Y (dependent variable) for each of the values of X (independent variable or predictor). In statistical jargon, this postulate is called "homoscedasticity." This postulate is impossible to respect with a binary dependent variable. Coded in this way, the average of the distribution is equal to the proportion of 1 in this distribution. For example, if we have a sample of 100 people and 30 of them are coded 1, the average of this distribution will be 0.30—that is, the proportion of 1. This average of 0.30 is also the probability of drawing a person randomly and this sample, which will be coded 1. The proportion and the probability of 1 are therefore the same in the current case. If we designate this proportion or probability with a P and the proportion of 0 with a Q in our sample ($Q = 1 - P$), the variance of such a distribution of PQ is 0.21. Therefore, when 50% is 1 ($P = 0.5$), the variance is 0.25, its maximum value. If $P = 0.10$, the variance becomes 0.09. The findings are the same if $P = 0.9$, so the closer P is to 1 or 0, the closer the variance gets to 0. This postulate cannot be respected.

2. The statistical test of B coefficients relies on the postulate that the residuals (Y-Y') are normally distributed. Since Y can only be 0 or 1, this postulate is very difficult to justify, even for approximations.

Linear regression is based on the least squares method, which strives to minimize the residual error term, while maximizing the variance due to the regression. A logistic analysis does not offer a mathematical solution for estimating the least squares method of parameters. Instead, the technique uses maximum likelihood, where likelihood represents a conditional probability P(Y|X), or the probability of Y for a given X. It is an iterative process that allows parameters to be extracted (a, b) that will be the most possible likelihood of this conditional probability P(Y|X). We will not go into the calculation in further detail, as that falls more within the domain of numerical analysis.

Probability and relative risk (odds)

We would like to illustrate the fact that probability and risk are not the same thing in logistic regression.

Take an example where 20 subjects (10 men and 10 women) are given a performance test for which success and failure are coded 1 and 0 respectively. We observe that the results for men are 7 successes and 3 failures, and that for women they are 3 successes and 7 failures. For men, the probability of success P will be 7/10 = 0.70, and the probability of failure Q will be 3/10 = 0.30. For women, P will be 3/10 = 0.30 and Q will be 7/10 = 0.70.

The risk (odds) of success is the odds ratio.
Odds for men = P/(1-P) = 0.7/(1-0.7) = 2.333
Odds for women = P/(1-P) = 0.3/(1-0.3) = 0.429

This is an asymmetrical situation, because the odds of success should be the opposite of the odds of failure. The natural logarithm allows us to transform everything so that the risk of one really is the opposite of the risk of the other. Thus, the ln(2.333.) is 0.847 and the ln(0.429) is -0.847. This is the desired opposite. Note that a risk of 0.5/0.5 transformed into a log becomes 0. We will see that the natural logarithms are the B coefficients of the logistic equation.

The odds ratio that the performance will equal 1 for men vs. women is 2.333/0.429 = 5.438. We can therefore say that the odds of success, i.e. the odds that the result of the performance is 1, are 5.438 times greater for men than for women.

In logistic regression, the dependent variable, called "logit," is the natural logarithm (ln) of the odds. The general formula is:
ln(odds) = logit(P) = ln(P/1-P) = ln(P/Q).

Hence, with a predictive variable, we will have logit(P) = a + bX. The logistic regression calculates the changes in the log odds of the dependent variable, not changes in the dependent variable itself, as is the case in normal linear regression.

Pseudo R-square

Logistic regression has no equivalent to the R-square of a standard multiple regression. Some mathematicians have examined the question, and the algorithms vary in such a way that the statistics often produce contradictory results. These pseudo R-squares do not signify the same thing as the proportion of variance of the dependent variable that can be explained by predictive variables, as in a normal regression. For this reason, we are not including this statistic in the report

Sensitivity and specificity

The independent variables of a model may be assessed using their sensitivity and specificity to describe a positive or negative binary phenomenon.
Take the case of the dependent variable *Sinusitis* with the following classification table:

		Cases of Sinusitis predictedusing logistic regression	
		Negative	Positive
Actual cases	Negative	79 (TN)	71 (FP)
of Sinusitis	Positive	10 (FN)	33 (TP)

Where TN = True negatives
TP = True positives
FP = False positives
FN = False negatives

Sensitivity is the ability to correctly identify cases that occur, i.e. true positives (TP).
The formula is: TP/(TP + FN) = 33/(33 + 10) = 76.7%.

Specificity is the ability to correctly identify cases that do not occur, i.e. true negatives (TN).
The formula is: TN/(TN + FP) = 79/(79 + 71) = 52.7%.

In practice, it is utopian to think that we could have both sensitivity and specificity equal to 100%.

The rate variable is the attack rate of symptoms in the sample being studied.
The formula is: (TP + TN) / (TP + TN + FP + FN).

Chi-square

Chi-square expresses the likelihood ratio between this model and an empty model, which is one with only the constant A of the equation, without coefficients. A significant chi-square indicates that the model as a whole is more reliable than an empty model.

Logit coefficient (logit)

These coefficients are also called unstandardized "logistic regression coefficients" and correspond to B coefficients in standard linear regression. These coefficients may be used to establish a prediction equation and thus generate predictive values, which, in this case, are logistic data.

In logistic regression, these B coefficients are the logits of the predictive variables involved in the equation used to estimate the ln(odds) that the dependent variable is equal to 1. The value of this logit represents the amplitude of the change in the ln(odds) of the dependent variable by unit of change (positive if the number is + or negative if the number is -) of the predictive variable.
In the preceding example, where we wanted to predict performance success on the basis of gender (where gender would be coded 1 for men and 0 for women), our B coefficient for the logistic equation would be 0.847, which does not reveal much more than the direction of the estimated change (positive weight between gender and success). This is why it is useful to transform this value into odds, using the formula exp(0.847) = 2.333.

Confidence intervals of 95% may be produced for each of the B coefficients of the logistic regression. If the value 1 is included between the upper and lower limits of this interval, this

means that this coefficient is not significantly different from 0 and brings nothing to the predictive model. In effect, if B=0, then exp(b) = 1, and a relative risk of 1 predicts nothing.

The "a" intercept in the equation describes a model that does not include any of the predictive variables, and, although it is reported here, is not very useful in the present context. It may eventually serve to generate logistic scores that may be helpful in other types of analysis.

Results of the logistic equations

We now present the coefficients and odds ratio of logistic equations. To help us properly interpret the reported values, we take the *Sinusitis* model as the dependent variable.

The logistic equation is:
Logit(*Sinusitis*) = -0.37 -1.70 *Any_Smoke* + 0.97 *Smoke.*

These are the coefficients reported in the first line of the variable in the following table.

Remember that these B coefficients indicate the quantity of the increase (positive number) or decrease (negative number) in the prediction that the risk log of *Sinusitis* is equal to 1 (attained) for each variation of 1 unit in the value of the predictor, when the predictive power of the other predictors is kept constant.

These coefficients are difficult to interpret because they are expressed in risk log units. This is why they are converted into relative risk exp(b) and are reported in the relative risk (often better known as the odds ratio) column. Another advantage of transforming these coefficients into relative risk is that it neutralizes the sign of this coefficient. Thus, an exp(3) = 20 (significant risk), while an exp(-3) = 0.05 (minimal risk).

Any_Smoke: For each increment of 1 (in this case, the fact of being 1 instead of 0), we estimate a decrease of 1.7 log (risk of *Sinusitis*), while keeping the influence of all other variables constant. The risk of getting it is exp(-1.70) = 0.18 times greater when compared to a zero contribution of this variable.

Smoke: For each increase of 1 in this variable, we observe an increase of 0.97 of log (risk of *Sinusitis*), while keeping the influence of all other variables constant. The risk of getting it is exp(0.97) = 2.63 time greater when compared to a zero contribution of this variable.