Credible Science, Influential Science?

Measuring the Credibility Attributed to Public Health Organizations in the Media

Antoine Lemor[1], Alexandre Fortier-Chouinard[2], François Claveau[1], Gabriel Saso-Baudaux[1], Francis Lareau[1], Adrien Cloutier[2], and Véronique Fournier[2]

[1]Université de Sherbrooke
[2]Université Laval

Working paper

Please do not cite or distribute without permission

Abstract

A public health crisis is not only biological but also inherently social. One of the clearest manifestations of the social dimension of health crises lies in the ways individuals access and evaluate information. For example, how audiences judge the credibility (i.e., whether information is believable, trustworthy, and expert) can shape behaviors and, ultimately, health outcomes. Yet the credibility attributed to public health organizations remains underexamined empirically and systematically. To address this gap, we develop a conceptual framework that links the two core dimensions of credibility, expertise and trustworthiness, to observable cues in news text, and implement an annotation protocol that combines expert review with supervised machine learning on an unprecedented corpus of 172,586 articles from more than 1,400 Canadian outlets spanning nearly five decades (since 1977). The study compares four organizations—the World health organization (WHO), Public health agency of Canada (PHAC), National public health institute of Quebec (INSPQ) and National institute for excellence in health and social services (INESSS). At the sentence level, we identify organizational mentions and messengers and tag positive or negative credibility cues alongside variables of interest (type of scientific evidence, public health domain, sentiments, etc.). This granular design yields a normalized credibility score at the article–organization level and, crucially, creates a direct bridge to analysis by allowing scores to be related to contextual features such as evidence type or domain. Beyond the indicator itself, this linkage enables systematic tests of the determinants of credibility in media coverage—for example, contrasting the effects of natural-science versus social-science evidence, or comparing domains such as mental health and epidemiology. The result is a reproducible tool to quantify, compare, and explain how media shape the perceived credibility of public health organizations.

1 Introduction

The credibility of an expert or an organization refers to "the likelihood of their claims being given weight by non-expert audiences" (Barnes 2005, p. 11). In public health, the credibility of organizations such as the World Health Organization (WHO) is central as it determines the extent to which their recommendations are perceived as reliable, valid, *and therefore influential*.

Studying the credibility of public health organizations becomes even more central when considering that public health crises are not solely biological but also social. For Horton (2020), who builds on anthropologist Merrill Singer's concept and definition of "syndemic" to describe the COVID-19 crisis, public health crises result from biological conditions but also social factors "that increase a person's susceptibility to harm or worsen their health outcomes."

One reason why individuals' health can be affected by social determinants lies in how they access information. The COVID-19 "infodemic", i.e. an "epidemic of information driven by the spread of rumors or false truths", is an exemple to the extent that it undermined individuals' health by disseminating misinformation, particularly about treatments or vaccines (proposed translation, Viallon, Dolbeau-Bandin, and Picot 2021, p. 10; Chen, Lee, and Lin 2022).

This study proposes a conceptual and methodological framework to measure the credibility attributed by the media to four organizations: the World Health Organization (WHO), the Public Health Agency of Canada (PHAC), the Public Health Agency of Canada (PHAC), the National Public Health Institute of Quebec (INSPQ), and the National Institute for Excellence in Health and Social Services (INESSS). The objective is to construct, based on an unprecedented corpus of over 170,000 Canadian media articles mentioning these organizations, a rigorous and reproducible measure of the credibility attributed to them.

Relying on machine learning methods, this study introduces a rigorous annotation protocol whose categories were developed in collaboration with four public health experts and practitioners from the National Public Health Institute of Quebec (INSPQ). These categories were designed to classify the sentences of each article and identify the credibility cues they contain. At this stage of the study, only the conceptual and methodological framework has been developed, and the results shown remain preliminary. The manual annotations will be carried out by a team of annotators, all co-authors of this study. Nevertheless, preliminary results were obtained by training models on annotations produced by generative LLMs, based on a methodology for which a technical paper will be released soon.

Preliminary results show that WHO shows negative average credibility, which likely reflects political controversies. Most notably, we observe a troubling inverse relationship between credibility and media prominence—articles with higher credibility scores are systematically less likely to receive front-page placement. This reveals a fundamental paradox where scientific credibility seems to reduce visibility. By cross-referencing this index with other media data (e.g., outlets, journalists, etc.) and additional variables (e.g., public health domains, types of scientific evidence, etc.), we aim to identify the factors that drive credibility.

2 Theoretical and conceptual framework

2.1 Credibility and influence of experts

The influence of experts refers to their ability "to shape policy decisions in line with their knowledge-based preferences," that is, in accordance with their recommendations (Christensen 2021, p. 463). However, this ability to influence largely depends on the credibility that experts and their organizations enjoy among the public and/or policymakers. The credibility of an expert or an organization refers to "the likelihood of [their] claims being given weight by non-expert audiences" (Barnes 2005, p. 11).

According to Viviani and Pasi (2017, p. 3), the literature agrees that credibility has two dimensions: *expertise* and *trustworthiness*. *Expertise* pertains to the intrinsic but perceived competencies and qualities of the information source. It helps assess the extent to which information appears valid and accurate. *Trustworthiness*, on the other hand, refers to the degree to which an audience perceives the communicator's claims as honest and believable, without any intent to deceive.

In public health, an organization deemed credible should be perceived as both *expert* and *trustworthy*. Specifically, being *expert* implies that the organization is believed to possess the knowledge, skills, resources, and experience necessary to assess a given situation. Being *trustworthy* means that the organization is expected to act honestly, transparently, and independently of any political or economic influence. The more credible an organization is perceived to be, the more likely it is to be influential; that is, to be heard and followed by the public and policymakers.

3 Methodology

3.1 Case studies and data

This study focuses on four organizations: (1) the World Health Organization (WHO), (2) the Public Health Agency of Canada (PHAC), (3) the National Public Health Institute of Quebec (INSPQ), and (4) the National Institute for Excellence in Health and Social Services (INESSS). Each of these organizations exhibits varying degrees of expertise in public health. Some are more "expert-oriented" (e.g., INSPQ), while others are more political (e.g., WHO) or administrative in nature (e.g., PHAC). Others specialize in services or medical technologies (e.g., INESSS). This study will explore these differences, particularly the hypothesis that organizations with a more political nature may be perceived as less credible. Articles mentioning these organizations were extracted from online databases (Factiva and Eureka) using a list of keywords detailed in Figure 1 below:

INSPQ

"INSPQ" or "Institut national de santé publique du Québec" or "Institut de santé publique du Québec" or "agence de santé publique du Québec" or "national institute of public health of quebec" or "quebec's national institute of public health" or "institute of public health of quebec"

INESSS

"institut national d'excellence en santé et services sociaux" or "INESSS" or "institut d'excellence en santé" or "institut national santé services sociaux" or "excellence santé et services sociaux" or "national institute for excellence in health and social services" or "Quebec health excellence institute" or "national institute Quebec health services"

ASPC

"agence de santé publique du canada" or "agence santé publique canada" or "ASPC" or "public health canada" or "PHAC" or "public health agency of Canada" or "public health agency canada" or "santé publique canada"

OMS

"organisation mondiale de la santé" or "OMS" or "agence sanitaire mondiale" or "organisation mondiale santé" or "règlement sanitaire international" or "world health organization" or "international health regulations"

Figure 1: Keywords used to extract media articles

These keywords were selected for their ability to capture all articles mentioning these organizations in both official languages of Canada (French and English). Once extracted, the data were filtered to retain only articles written by journalists or opinion pieces and letters. The resulting database comprises a total of 172,586 articles (Table 1) from 1,483 media outlets. The top 50 media outlets represented in the database are shown in Figure 2.

Table 1: Distribution of media articles by organization

INESSS	INSPQ	ASPC	OMS	**Total**
4 770	29 372	44 208	94 236	**172 586**

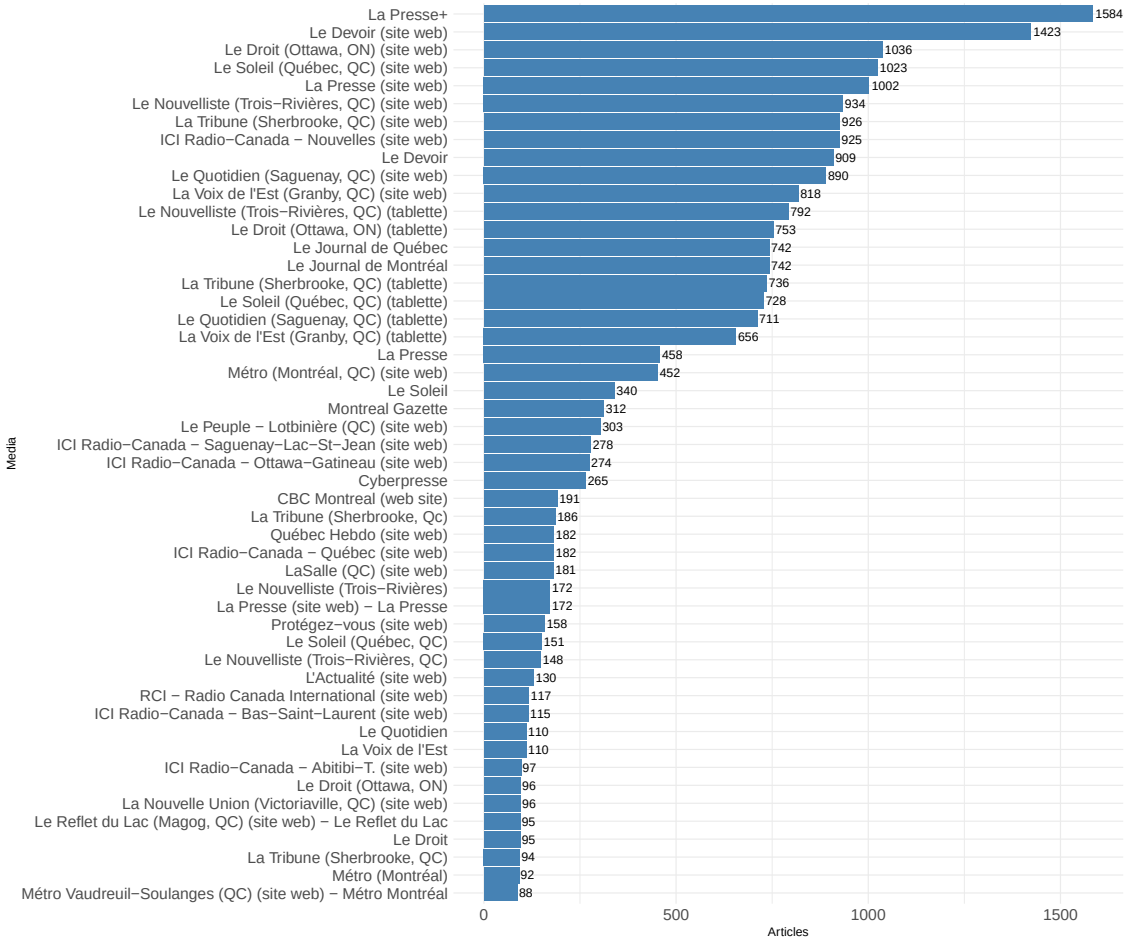


Figure 2: Distribution of articles by media outlets (top 50)

3.2 Text analysis methodology

3.2.1 Machine learning

This study relies on machine learning to classify each sentence of every article based on the presence of specific categories (Do, Ollion, and Shen 2022). Machine learning is a data analysis approach that enables an algorithm to learn from training data. For this study, we

use supervised classification models to identify entities and credibility cues within articles, using a set of manual annotations performed by a team of annotators, all co-authors of this study. We also test a fork of the library build by Do, Ollion, and Shen (2022) we improved for social science studies and scientific research (Lemor 2025).

3.2.2 Modeling the information communication process in articles

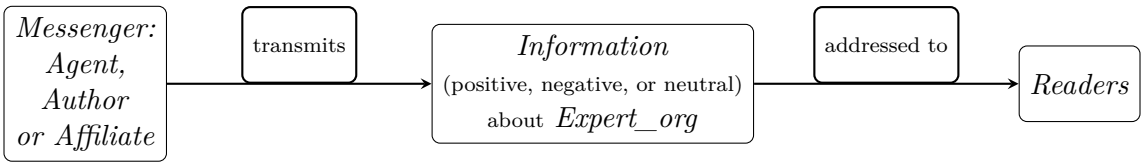
At the sentence level (which is the focus for the annotation process we aim to implement), we model the communication process as follows:

> a positive, negative, or neutral information (denoted as *Information*) about one of the four organizations (denoted as *Expert_org*) reported by a journalist is transmitted by a messenger to the audience (*Readers*).

The *Messenger* can be (1) the author of the article itself, for example in the case of an opinion piece (denoted as *Author*), (2) an external agent whose message is relayed by the author (denoted as *Agent*), or an affiliate of one of the four organizations cited by the author (denoted as *Affiliate*). Figure 3 illustrates this process.

The manual and automated annotation will aim to identify in each sentence all the categories presented in Figure 3: the *Information*, the *Expert_org*, and the various *Messengers*, as well as *other variables of interest contained within the Information* (notably: evidence from natural sciences, evidence from social sciences, public health domains, sentiments, etc.).

Figure 3: Information communication schema at the sentence level: *Information* about an *Expert_org* is transmitted by a *Messenger* (*Agent*/*Author*/*Affiliate*) to the *Readers*.



3.2.3 Measuring credibility: credibility cues

The content of the *Information* transmitted, as illustrated in Figure 3 above, is likely to contain cues that help assess the credibility of the selected organizations. In other words, the transmitted *Information* includes elements that influence how *Readers* perceive the credibility of organizations or experts. Consider the following example taken from one of the articles in the database:

> "A lawyer leading the class action lawsuit against the Government of Canada regarding the listeriosis outbreak in the country, Mr. Evatt Merchant, *criticizes the lack of information made public about the confirmed listeriosis case in New Brunswick*. Mr. Merchant *finds it 'odd' that the officials from the Public Health Agency of Canada disclose so little information about this listeriosis case.*"

Cues	Credibility Dimension	Definition	Examples
Positive credibility cues			
Precise scientific results or quantified claims	Expertise	This category encompasses all sentences explicitly presenting: **quantified results**, **precise statistics**, or **specific conclusions** drawn directly from an **identifiable research study**; **recognized scientific claims** that are accepted without apparent debate or uncontested within the **scientific community**, and generally expressed by a **recognized organization** whose **scientific authority** is used as an argument to reinforce the credibility of the statement.	"This represents a 'very significant' increase of 85.7% compared to the same period last year"
Association with socially valued objectives or inter-institutional collaboration	Trustworthiness	Describes any situation where **research**, a **scientific study**, or an **organization's action** (recommendation) is explicitly linked to a **concrete objective of social improvement**, whether practical or in terms of public health, safety, environment, or economy. Emphasis is placed both on the final practical or societal goal and on **cooperation between multiple actors** (decision-makers, public agencies, expert institutions, etc.) to achieve this objective.	"Universal and low-cost childcare services for parents are essential to women's participation in the labor market, according to INSPQ"
Indication of titles or status	Expertise and Trustworthiness	Any explicit mention of **academic titles**, **positions**, or **professional status** of individuals involved in research (professors, doctors, principal researchers, etc.) or within the organization and who are cited in the article. This allows recognition of the authority or expertise of the cited individual.	"Dr. Généreux is currently working on writing her scientific article for imminent publication"
Information about scientific publication	Expertise	Specific mention of an ongoing **scientific publication** (article, journal, conference, preprint, etc.) or the **scientific platform** where the research has been published. This category is used exclusively for precise identification of a scientific publication.	"Dr. Généreux is currently working on writing her scientific article for imminent publication"
Preparedness and resources	Expertise	Specifically concerns **preventive actions**, **strategies**, or **proactive measures** implemented to anticipate or manage a future crisis (health, environmental, security, etc.). This includes **strategic plans**, **medical equipment reserves**, **drug or vaccine procurement**, **establishment of expert units**, etc., and any **additional resources** deemed scalable to improve the organization's action.	"The spokesperson for the Public Health Agency of Canada (PHAC), Michael Andrew Tansey, states that the federal agency periodically updates its pandemic influenza response plan."
Transparent communication or valorization of organizational independence	Trustworthiness	Refers to any situation in which a **scientific organization** highlights, on one hand, its commitment to **transparency** (explicit willingness to make data, methods, and results understandable and accessible to all stakeholders, including the general public) and, on the other hand, its **independence**. This independence can be emphasized by asserting **autonomy** from **political, financial**, or other influences, with the aim of guaranteeing the objectivity and credibility of its work.	"There is hardly any precedent where a Health Minister has overturned a decision by the Institute, an independent organization."
Negative credibility cues			
Methodological weaknesses, missing data, or uncertainty	Expertise	Refers to any mention of **limitations**, discrepancies, or questioning regarding the **research methodology** employed. This may involve unusual approaches, insufficiently robust tools, partial analyses, or inadequate validation processes. Also refers to any situation where **cited data are incomplete, inconsistent, or uncertain**, or when the lack, revision, or limited reliability of certain data is explicitly highlighted.	"We're going a bit against what is normally done in research, but for me the most important thing is that the population is informed as well as decision-makers," explains the researcher (...)"
Politicization	Trustworthiness	This category exclusively concerns cases where a **scientific organization** (e.g., INSPQ, INESSS, WHO) explicitly takes a position or proposes a **public policy option** without providing **scientific justification**, and this absence of scientific justification is explicitly highlighted in the article. It also includes cases where the organization's position is explicitly described as political in nature in the text, or when a **political decision-maker** directly and explicitly intervenes regarding a recommendation or decision by the scientific organization.	"Health Minister Yves Bolduc forced INESSS to review its decision by approving certain expensive cancer treatment drugs that it had refused."
Controversial position or scientific dissensus	Expertise and Trustworthiness	Presence of a **significant debate** or **manifest contradiction** regarding commonly accepted scientific norms, standards, or conclusions. This category highlights **notable disagreement** within the **scientific community**, whether arising from a minority position against an established standard or from divergence between multiple studies on the same subject.	"In 2008, Health Canada lowered the recommended maximum threshold from 800 Bq/m^3 to 200 Bq/m^3, twice the reference level set by WHO."
Organizational failure and unpreparedness	Expertise	Refers to any situation where the expert organization fails in its capacity to **anticipate** or adequately fulfill its responsibilities. This failure may manifest as an **absence or insufficiency of planning** or as an **operational shortcoming** (protocol adherence, communication of critical information, etc.).	"Hogan said that was a failure of the agency to come up with a common plan and something they should address."
Lack or reduction of resources	Expertise	Refers to any situation where **resources** (**human, financial, material**, etc.) are **reduced, insufficient**, or not allocated adequately to meet the needs of a project or policy related to the expert organization.	"Some 2,224 positions are set to disappear. In total, Quebec risks losing 236 jobs, Ontario 432, and the Prairies 775."
Lack of transparency or questioning of independence	Trustworthiness	Explicit criticism regarding the organization's **transparency**, particularly the difficulty in obtaining **precise data**, and questioning of its **independence** or **objectivity**.	"Lawyer Evatt Merchant criticizes the little information made public about the confirmed case of listeriosis in New Brunswick."

Table 2: Taxonomy of positive and negative credibility cues.

In this case, we observe that a *Messenger* (Mr. Evatt Merchant) conveys negative *Information* about an *Expert_org* (PHAC). This negative *Information* undermines one of the two dimensions of credibility discussed in section 2.1: the organization's *trustworthiness*. Indeed, the alleged lack of transparency calls into question the organization's perceived intentions, and therefore its *trustworthiness*. In other words, the absence of transparency is perceived here as a *negative credibility cue*.

Through careful reading of articles randomly selected from the database, several types of credibility cues were identified by the first author: *negative cues* and *positive cues* of credibility. These cues were constructed and classified according to the two dimensions outlined in section 2.1: *expertise* and *trustworthiness*. For each cue, mutually exclusive definitions were proposed, discussed and updated with four public health experts from INSPQ to ensure their relevance. Table 2 above presents the list of credibility cues retained through this process.

3.2.4 Organizations, messengers and variables of interest

In addition to credibility cues, our methodology aims to identify the *Expert_org*, *Messengers*, as well as other *Variables of interest*. The objective is to subsequently cross-reference these variables with one another in order to: (1) measure the credibility of organizations, (2) measure credibility by category of messengers, and (3) measure the impact of variables of interest on the credibility of organizations. For each sentence, in addition to credibility cues, we therefore identify the following *Expert_org* (Table 3):

Table 3: Taxonomy of *Expert_org* used in the study

Category	**Definition**
INSPQ	Any reference to INSPQ.
WHO	Any reference to WHO.
INESSS	Any reference to INESSS.
PHAC	Any reference to PHAC.

We also identify the following *Messengers* (Table 4):

Table 4: Taxonomy of *Messengers* used in the study

Category	**Definition**
Public health authorities (DNSP, DRSP, etc.) other than public health organizations	Refers to any governmental instance, organization, or service specialized in surveillance, prevention, and protection of population health but which fall under federal or provincial administration and are neither INSPQ, INESSS, nor PHAC, nor political decision-makers.
Other administrative authorities	Includes Canadian or Quebec public agencies, departments, or entities that do not directly fall under public health but exercise various administrative, regulatory, or operational functions (e.g., customs, border services agencies, ministries of Education, Environment, etc.).
Political decision-makers	Refers to any person or entity invested with formal authority to develop, direct, or influence public policies exclusively elected officials (ministers, parliamentarians, mayors, etc.).
Public, patients, and population	Groups all references, mentions, or descriptions that explicitly concern individuals directly affected by a public health issue or by one of the expert organization's missions, whether it be the general public as a whole, particular populations (e.g., at-risk groups, local communities), or patients receiving care.
Stakeholders	Any actor involved or having a direct or indirect interest in a policy, measure, or public health situation related to the expert organization's missions. This includes, but is not limited to, health professionals (excluding patients), non-governmental organizations, professional associations, businesses, unions, interest groups, and international (e.g., WHO, UN) or community organizations.

We also identify the following *Variables of interest* (Table 5):

Table 5: Taxonomy of *Variables of interest* used in the study

Category	Definition
Natural sciences evidence	Encompasses any evidence, data, or results based on methods or concepts from natural sciences (biology, epidemiology, chemistry, physics, etc.). This includes reproducible methodologies and data on virus evolution, epidemiological statistical analyses, environmental measurements, etc.
Social sciences evidence	Concerns evidence or data drawn from social science disciplines (sociology, psychology, economics, political science, etc.). This includes studies on human behavior, perception surveys, field survey data on social impacts, etc.
Public policy recommendations or proposals	Refers to any suggestion, proposal, or official advice concerning the adoption, revision, or implementation of public policies or specific measures. This covers both expert recommendations and those mentioned on behalf of expert organizations.
Government criticism	Groups all statements, viewpoints, or unfavorable judgments that question, contest, or criticize the choices, actions, or omissions of a government or public authority.
Public health domains	1. COVID-19; 2. Infectious diseases and epidemiology; 3. Environmental health and food safety; 4. Mental health and addictions; 5. Maternal, child, reproductive, and sexual health; 6. Specific populations (workplaces, elderly, etc.) or vulnerable groups; 7. Health promotion, education, equity, or health inequalities; 8. Medical technologies, medical research, and innovation; 9. Health system policies, legislation, and organization; 10. Organization and funding of INSPQ/INESSS/WHO/PHAC; 11. Surveillance and monitoring.
Public policy domains	Categorization from the *Comparative Agendas Project*: 1. Justice and crime; 2. Immigration; 3. Science, technology, and communications; 4. Economy and finance; 5. Employment and labor; 6. Transportation and infrastructure; 7. Housing, urban and rural development; 8. Natural resources, agriculture, environment, and energy; 9. International affairs and defense; 10. Governance and public institutions; 11. Culture, identity, and nationalism; 12. Rights, freedoms, education, health, and social protection.
Positive sentiments	Sentences in which we detect a connotation of optimism, relief, confidence, or satisfaction related to the described situation (e.g., a public health action, a scientific discovery, the effectiveness of a measure), or a favorable evaluation of an actor (organization, expert, government, etc.).
Negative sentiments	Sentences conveying an unfavorable emotion or pejorative assessment (fear, concern, anger, distrust, disappointment, frustration, etc.). These emotions may relate to the health situation, to an actor's response (e.g., expert organizations), or to any other described reality.

3.2.5 Annotation Process

For each element to be annotated (*Expert_org*, *Messenger*, *Credibility cues*, *Variables of interest*), an annotation process will be implemented with 6 annotators, and one project lead. This process will rely on careful reading of sentences randomly selected from the database, followed by manual annotation aimed at identifying each of the abovementioned categories. The manual annotations serve to (1) create a training dataset for training BERT-type models and (2) validate the effectiveness of the training with a test dataset. Inter-annotator comparison measures will be implemented to assess the reliability of annotation (Krippendorff's alpha), and the consistency of categories (Krippendorff n.d.). Let us revisit the following example:

> "A lawyer leading the class action lawsuit against the Government of Canada regarding the listeriosis outbreak in the country, Mr. Evatt Merchant, criticizes the lack of information made public about the confirmed listeriosis case in New Brunswick. Mr. Merchant finds it 'odd' that the officials from the Public Health Agency of Canada disclose so little information about this listeriosis case."

Annotation will be performed at the sentence level, but reading is conducted on the entirety of each randomly selected article to ensure complete context and better annotation quality. Based on the previous example, an annotator should assign the following labels:

– Negative credibility cue: *Lack of transparency or questioning of independence*

– Expert_org: *PHAC*

– Messenger: *Public, patients, and population; Political decision-makers*

– Variables of interest: *Government criticism; Negative sentiments; Environmental health and food safety*

3.2.6 Proposed credibility index

This working paper is limited to describing the annotation procedure. However, one can, for example, construct the final credibility index according to the following lines:

– Each positive index (reinforcing credibility) is worth +1.

– Each negative index (undermining credibility) is worth −1.

– Each neutral or ambiguous index is evaluated at 0.

The *credibility score* for an article can thus be expressed by *relating the sum of cues to the length of the text*, for example to the total number of sentences:

$$\text{Score}_{\text{article}} = \frac{(\#\text{positive}) - (\#\text{negative})}{\text{Total number of sentences}}.$$

The "positive" and "negative" correspond to the occurrences annotated in the article according to the stated rules, while neutral occurrences do not modify the score. The indices thus obtained can subsequently be aggregated by organizations or by any other desired level of analysis.

For example, for a given organization, one can calculate a *global credibility index*:

$$\text{Index}_{\text{organization}} = \frac{\sum_{i=1}^{N} \left(\text{Score}_{\text{article}_i}\right)}{N},$$

where N is the total number of articles concerning the organization. This yields a global measure of the credibility attributed to the given organization. The entire annotation process is described in Figure 4.

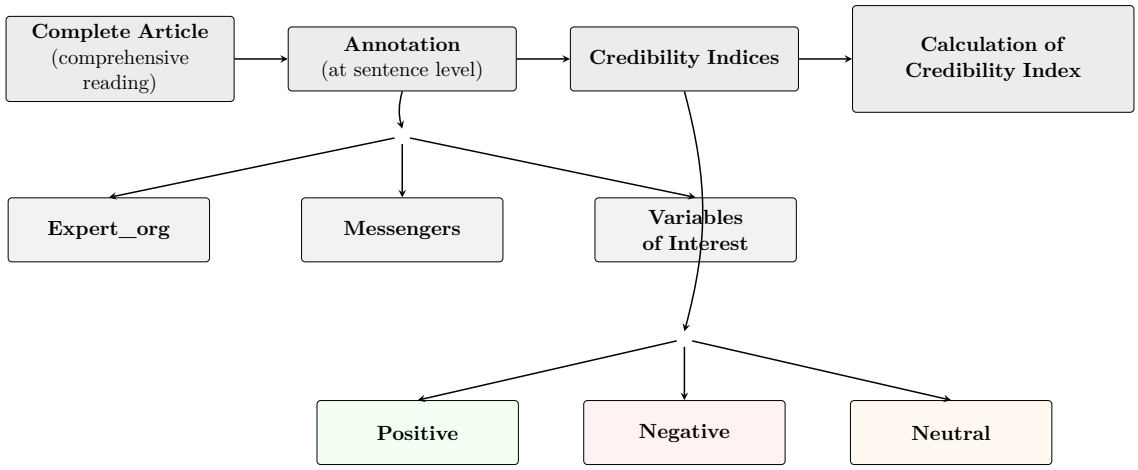


Figure 4: Schema describing the annotation process and construction of a credibility index.

4 Preliminary results

4.1 Implementation of an automated annotation framework

Following the development of our conceptual framework, we implemented an automated annotation process using generative LLMs and supervised machine learning models to produce preliminary results for this conference presentation. We generated 3,000 manual annotations based on prompts that captured the credibility cues outlined in our taxonomy.

These annotations served as training and test data for training BERT-type models (Camembert in French), which we then used to classify the remaining corpus of French-language articles according to our predefined categories of positive and negative credibility cues, organizational mentions, and variables of interest (type of evidence, public health domains, etc.). For training the Cambembert models, we used a 80/20 split for training and a modified version of the library originally developed by Do, Ollion, and Shen (2022), to which we added new functionalities (Lemor 2025).

It is important to note that this automated approach was implemented specifically to produce preliminary findings for conference discussion. As outlined in our methodology section, the full implementation of this study will rely on comprehensive manual annotations by our team of annotators. The results presented here therefore do not include F1 scores or other validation metrics that will be essential for the final study.

The automated annotation process nevertheless successfully analyzed the French subset of our corpus. It allowed for the construction of preliminary normalized credibility scores for each of the four public health organizations under study. We focused exclusively on French-language articles for this preliminary analysis to facilitate more manageable data processing for conference presentation purposes, as the French subset represents a smaller portion of our total corpus compared to the complete bilingual dataset.

4.2 Analytical framework and statistical models

Our analysis employs three complementary approaches to measure credibility. First, we construct a normalized credibility index at the article level, calculated as the difference between positive and negative credibility cues divided by the total number of sentences in each article. This normalization ensures comparability across articles of varying lengths and prevents longer articles from mechanically accumulating higher absolute credibility scores.

Second, we implement an attribution algorithm that assigns credibility cues to specific organizations mentioned in the text. Given that credibility cues and organizational mentions may not co-occur in the same sentence, we developed an "outward search" algorithm that identifies the nearest organizational mention to each credibility cue, searching progressively outward from the tagged sentence until an organization is found.

Third, we estimate two complementary statistical models using the aggregated article-level data:

(1) We regress the normalized credibility index on a set of binary predictors including public health domains (e.g., COVID-19, infectious diseases, mental health) and other variables of interest mentioned in subsection 3.2.4 (e.g., presence of natural science evidence,

positive sentiment, government criticism). This OLS model identifies which factors are associated with higher or lower credibility scores.

(2) We model the probability of first-page placement as a function of the same predictors, augmented with two binary indicators derived from the credibility index: whether the article has positive credibility (index > 0) and negative credibility (index < 0). This specification allows us to test whether credibility itself influences media prominence while controlling for content characteristics. Page number metadata are available for many articles in our corpus which enables this analysis.

The database architecture supports these analyses through segmented tables containing sentence-level annotations linked to article-level metadata (we used Spacy models for tokenizers). Each sentence is tagged with binary indicators for credibility cues, public health domains, and other variables of interest. We also used NER (Named Entity Recognition) models to identify organizational mentions that were used in our attribution algorithm mentioned above.

4.3 Organizational credibility patterns

Our analysis show distinct credibility profiles (Figure 5). INSPQ emerges as the organization with the highest average credibility score (approximately 0.42), followed by PHAC (0.22), while INESSS shows a near-neutral score (0.01), and WHO displays a slightly negative average score (-0.04). These differences are statistically significant, as indicated by the non-overlapping 95% confidence intervals (except between the PHAC and the INESSS).

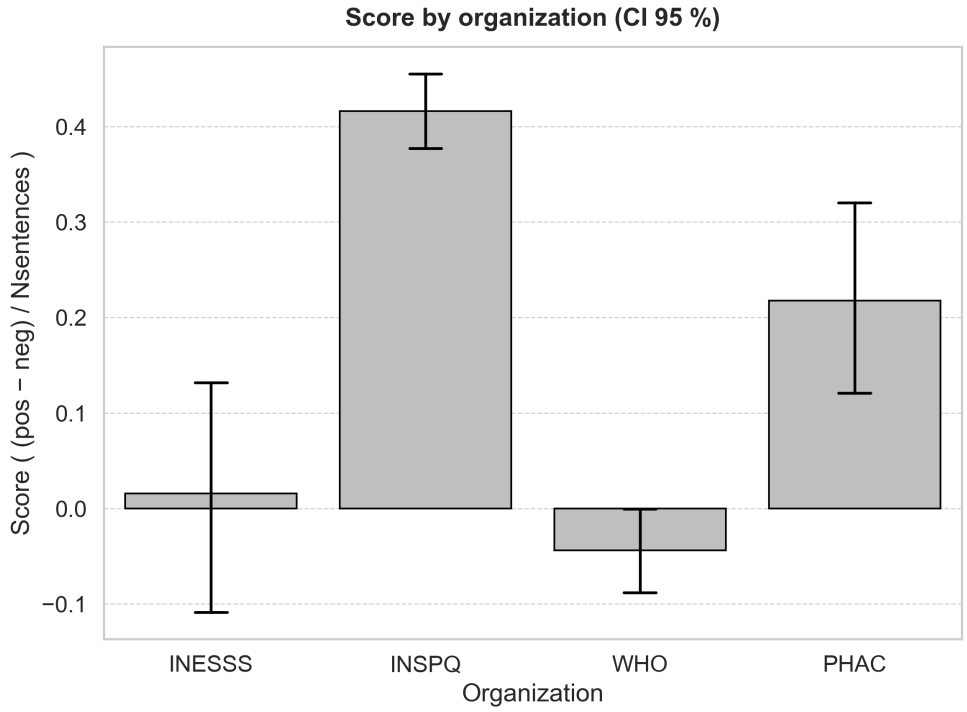


Figure 5: Credibility scores by organization with 95% confidence intervals

The superior credibility attributed to INSPQ in Canadian French-language media aligns with our theoretical expectation that more specialized, expert-oriented organizations may enjoy greater perceived credibility compared to those with more political or administrative functions. The negative score for WHO, despite its international scientific authority, suggests that its political dimensions and global controversies may influence how it is portrayed in Canadian media discourse.

4.4 Temporal dynamics of credibility

The longitudinal analysis of INSPQ's credibility over nearly two decades reveals significant temporal variation (Figure 6). Most notably, during the COVID-19 pandemic period (2020-2023), INSPQ enjoyed substantially positive credibility scores. This sustained credibility suggests that the organization's scientific expertise was particularly valued.

The 3-month moving average trend line shows this sustained positive credibility throughout most of the pandemic period. These temporal patterns indicate that organizational credibility is highly responsive to contextual factors, and further qualitative analyses will be conducted to better understand the specific determinants of these fluctuations.

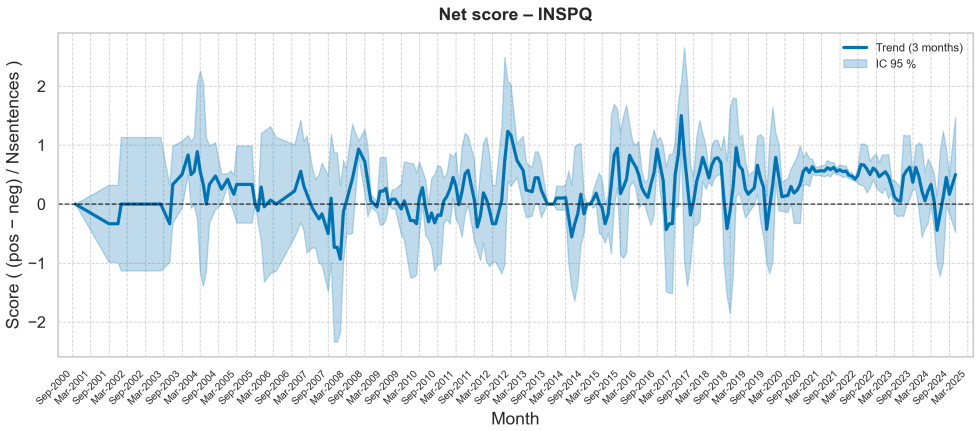


Figure 6: Temporal evolution of INSPQ's credibility score with 3-month moving average and 95% confidence intervals

In contrast, the World Health Organization's credibility presents a different pattern (Figure 7). WHO's credibility shows considerably more volatility over time. However, as the preliminary results were produced only on French-language articles, the number of observations decreases considerably which makes the results difficult to interpret (the confidence intervals are very wide).

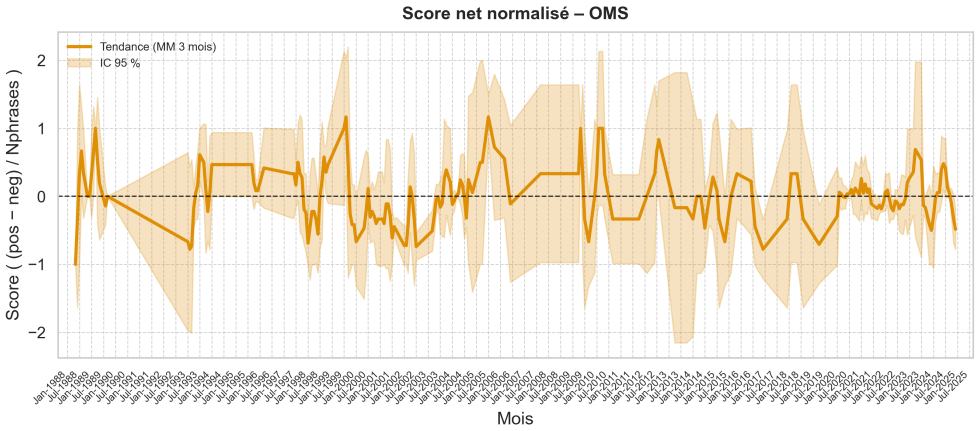


Figure 7: Temporal evolution of WHO's credibility score with 3-month moving average and 95% confidence intervals

4.5 Determinants of credibility and media prominence

Our models provide insights into both the factors that influence credibility scores and those that affect an organization's likelihood of receiving prominent media coverage (first-page placement).

4.5.1 Effects on credibility scores

For INSPQ specifically, several variables demonstrate significant associations with credibility scores (Figure 8). The most notable finding is that government criticism shows the strongest negative association with credibility, followed by negative sentiment. This shows that political controversies can substantially reduce perceived credibility.

Among public health domains, health systems and policies show one the strongest positive association with credibility, followed by content addressing vulnerable populations and infectious diseases. This pattern likely reflects INSPQ's established expertise in epidemiological monitoring and outbreak response. Health promotion and equity topics also demonstrate positive effects, as do recommendations, which shows that when INSPQ makes a recommendation, it tends to enhance the organization's perceived credibility.

Interestingly, the presence of social science evidence shows a moderate negative association with credibility, while natural science evidence has a small positive effect. This suggests that audiences may perceive INSPQ's credibility differently depending on the type of scientific evidence presented, with social sciences potentially viewed as less authoritative. Food and environmental health also show negative associations, possibly indicating areas where INSPQ's expertise is less recognized or where public skepticism is higher.

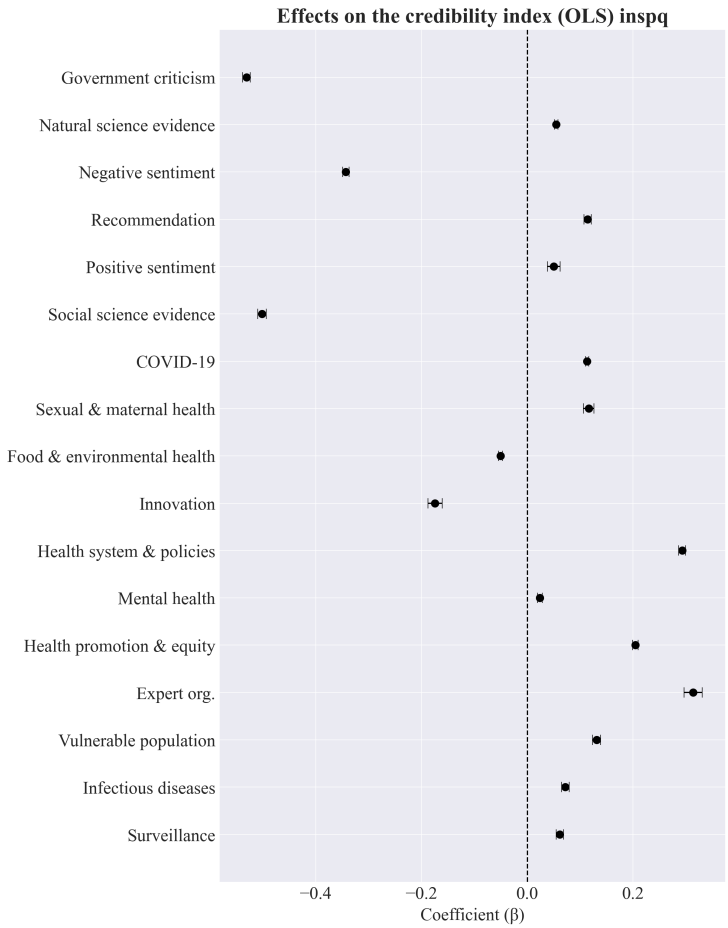


Figure 8: Forest plot showing effects of various factors on INSPQ's credibility index (OLS coefficients)

4.5.2 Effects on Media Prominence

The logistic regression analysis examining first-page probability reveals distinct patterns (Figure 9). Negative sentiment shows the strongest positive effect on front-page placement ($\beta \approx 0.85$), which is well known: critical or controversial content is substantially more likely to receive prominent media positioning.

This pattern is reinforced by government criticism ($\beta \approx 0.40$) and natural science evidence ($\beta \approx 0.45$), both of which increase the likelihood of first-page coverage.

Among public health domains, health system and policies ($\beta \approx 0.30$), innovation ($\beta \approx 0.25$), and positive sentiment ($\beta \approx 0.20$) show moderate positive effects on front-page probability. Food and environmental health topics also increase prominence likelihood ($\beta \approx 0.15$), as do recommendations ($\beta \approx 0.10$) and vulnerable population coverage ($\beta \approx 0.10$).

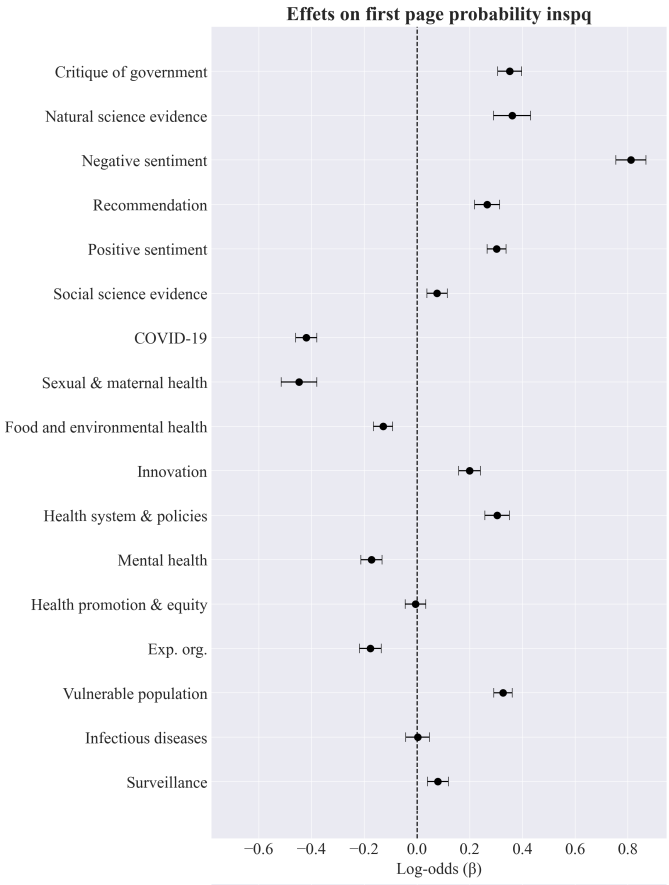


Figure 9: Forest plot showing effects on first-page probability for INSPQ articles (log-odds coefficients)

Several domains show minimal effects on media prominence, including surveillance, infectious diseases, health promotion and equity, expert organizations, and mental health (all with β values near zero). Sexual and maternal health ($\beta \approx -0.40$) and COVID-19 content ($\beta \approx -0.35$) show negative associations with front-page placement, possibly reflecting media fatigue with pandemic coverage or the routine nature of sexual health reporting.

The presence of social science evidence shows a slight negative effect ($\beta \approx -0.05$), contrasting with the positive effect of natural science evidence. This differential treatment of evidence types in media prominence decisions suggests that journalists may perceive natural science as more newsworthy or authoritative than social science.

4.6 The credibility-prominence paradox

A particularly striking finding emerges when examining the direct relationship between credibility scores and media prominence. Figure 10 illustrates the predicted probability of first-page placement as a function of the credibility index across all organizations. The analysis reveals a clear negative relationship: articles with higher credibility scores are systematically less likely to receive prominent media placement.

This relationship represents a fundamental paradox in public health communication:

the more credible an organization appears in media discourse, the less likely its coverage is to receive prominent placement. Articles with the lowest credibility scores (around -3) have approximately 8.6% probability of first-page placement, while those with the highest credibility scores (around +3) have only about 5.7% probability—a reduction of nearly one-third.

This suggests that media logic may systematically prioritize conflict, controversy, and criticism over positive or neutral scientific content. From a public health perspective, this creates a perverse incentive structure where organizations seeking media attention might benefit from generating controversy, while those maintaining high scientific standards and avoiding conflict receive less visibility. This finding has profound implications for how public health organizations should balance credibility maintenance with communication effectiveness.

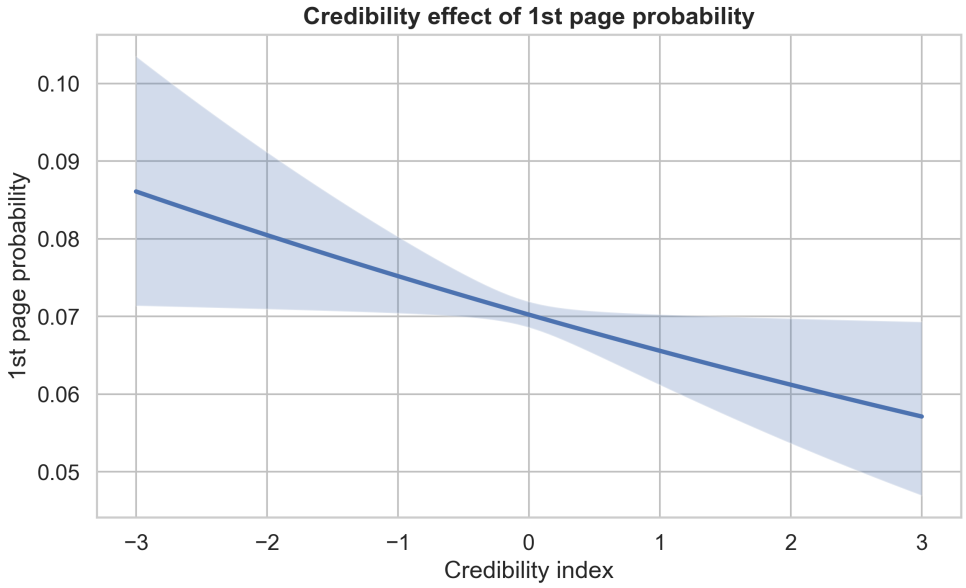


Figure 10: Predicted probability of first-page placement as a function of credibility index (global analysis)

4.7 Evidence types and credibility

Our analysis also reveals differential effects of natural science versus social science evidence on credibility attribution. Natural science evidence consistently shows positive associations with credibility scores, while social science evidence shows more modest positive effects. This pattern suggests that media discourse may privilege certain types of scientific evidence, potentially reflecting broader societal hierarchies of scientific knowledge. The relationship between evidence types and credibility provides empirical support for our theoretical framework's emphasis on expertise as a key dimension of credibility. Organizations are perceived as more credible when their claims are grounded in recognized scientific evidence, particularly from the natural sciences.

4.8 Implications for public health communication

These preliminary findings have important implications for understanding how public health organizations can maintain and enhance their credibility in media discourse. The differential credibility profiles suggest that organizational characteristics—degree of specialization, political independence, and mandate focus—significantly influence media portrayal. The temporal variation in credibility scores indicates that organizations must continuously manage their reputational capital, as credibility can fluctuate substantially over time.

The credibility-prominence paradox identified in our analysis presents perhaps the most significant challenge for public health communication strategy. Organizations face a fundamental trade-off between maintaining high credibility through rigorous scientific standards and achieving the media visibility necessary for public influence. This finding suggests that the media ecosystem may systematically undervalue credible scientific communication while amplifying controversial or critical content.

The distinction between factors affecting credibility scores and those influencing media prominence highlights the complexity of media dynamics in public health communication. Organizations must deal with multiple and even conflicting objectives: maintaining scientific rigor, achieving public visibility, and preserving institutional credibility over time.

These results provide a foundation for more detailed analyses of credibility determinants and offer practical insights for public health organizations seeking to optimize their communication strategies while preserving public trust. Future research should explore strategies that allow organizations to maintain both credibility and visibility, potentially through targeted communication approaches that recognize and work within existing media incentive structures.

5 Conclusion

This study advances our understanding of credibility by developing and implementing a comprehensive framework to measure the credibility attributed to public health organizations in Canadian media. Through the analysis of an unprecedented corpus of 172,586 articles from 1,483 Canadian media outlets spanning nearly five decades, we demonstrate both the feasibility and importance of systematically quantifying organizational credibility.

Our preliminary findings reveal striking patterns that challenge conventional assumptions about the relationship between scientific expertise and media influence. The hierarchical credibility ordering, with INSPQ enjoying the highest credibility scores, followed by PHAC, INESSS, and WHO, suggests that organizational characteristics matter profoundly. More specialized, expert-oriented organizations like INSPQ appear to benefit from greater perceived credibility compared to those with broader political mandates like WHO, whose negative average score likely reflects the political controversies surrounding its global role. The temporal analysis uncovers significant volatility in organizational credibility, with INSPQ experiencing a substantial credibility during the COVID-19 pandemic.

Perhaps our most consequential finding is the credibility-prominence paradox: articles with higher credibility scores are systematically less likely to receive front-page placement. This inverse relationship between scientific credibility and media visibility represents a fundamental challenge for public health communication. Organizations face an impossible choice

between maintaining rigorous scientific standards that enhance credibility but reduce visibility, or engaging in controversy that attracts media attention but erodes public trust. This structural bias in media systems toward conflict and controversy over scientific consensus has profound implications for public health influence.

The differential effects of evidence types further illusrates the media's hierarchical treatment of scientific knowledge. Natural science evidence enhances both credibility and media prominence, while social science evidence shows negative associations with credibility. This finding suggests that media discourse continues to privilege certain forms of scientific authority, potentially limiting the influence of behavioral and social approaches to public health challenges.

Our methodological contributions extend beyond these empirical findings. The development of mutually exclusive credibility cue categories through collaboration with INSPQ practitioners ensures that our measurement framework captures dimensions of credibility that matter in practice. The combination of theoretical grounding in the expertise-trustworthiness dyad in credibility with empirical validation through machine learning provides a reproducible approach that can be extended to other contexts and organizations.

Several limitations warrant acknowledgment. Our preliminary analysis focuses exclusively on French-language media, and the automated annotation process used for the APSA conference presentation will be replaced by comprehensive manual annotation in the full study. The attribution algorithm's reliance on proximity assumptions may miss more complex discursive relationships between credibility cues and organizational mentions and must also be validated. Future iterations should also explore weighted credibility indices that account for the differential importance of various cue types.

Looking forward, this research establishes the foundation for a systematic observatory of public health credibility. The ability to track credibility dynamics over time and across organizations provides essential intelligence for evidence-based communication strategies. Organizations can identify which domains enhance or undermine their credibility, understand how different types of scientific evidence are received, and anticipate the media dynamics that shape public discourse.

As our findings demonstrate, credibility directly influences media prominence and, by extension, public influence. In an era of increasing skepticism toward scientific institutions and proliferating misinformation, understanding and managing organizational credibility becomes essential for effective public health action. The tools developed here offer a path toward more strategic, evidence-based approaches to building and maintaining the public trust upon which all public health interventions ultimately depend.

6 References

Barnes, Barry (Mar. 2005). "The Credibility of Scientific Expertise in a Culture of Suspicion". In: *Interdisciplinary Science Reviews* 30.1, pp. 11–18. ISSN: 0308-0188. DOI: 10.1179/030801805X25882. (Visited on 09/12/2023).

Chen, Xi, Woohyung Lee, and Fen Lin (Jan. 2022). "Infodemic, Institutional Trust, and COVID-19 Vaccine Hesitancy: A Cross-National Survey". In: *International Journal of Environmental Research and Public Health* 19.13, p. 8033. ISSN: 1660-4601. DOI: 10.3390/ijerph19138033. (Visited on 04/17/2025).

Christensen, Johan (July 2021). "Expert Knowledge and Policymaking: A Multi-Disciplinary Research Agenda". In: *Policy & Politics* 49.3, pp. 455–471. DOI: 10.1332/030557320X15898190680037.

Do, Salomé, Étienne Ollion, and Rubing Shen (Dec. 2022). "The Augmented Social Scientist: Using Sequential Transfer Learning to Annotate Millions of Texts with Human-Level Accuracy". In: *Sociological Methods & Research*, p. 00491241221134526. ISSN: 0049-1241. DOI: 10.1177/00491241221134526. (Visited on 04/15/2024).

Horton, Richard (Sept. 2020). "Offline: COVID-19 Is Not a Pandemic". In: *The Lancet* 396.10255, p. 874. ISSN: 0140-6736, 1474-547X. DOI: 10.1016/S0140-6736(20)32000-6. (Visited on 07/18/2024).

Krippendorff, Klaus (n.d.). "Computing Krippendorff's Alpha-Reliability". In: ().

Lemor, Antoine (June 2025). *Antoinelemor/AugmentedSocialScientistFork*. (Visited on 06/19/2025).

Viallon, Philippe, Cécile Dolbeau-Bandin, and Jérémy Picot (2021). "Introduction:L'infodémie entre information et désinformation". In: *Les Cahiers du numérique* 17.3, pp. 9–15. ISSN: 1622-1494. DOI: 10.3166/LCN.2021.010. (Visited on 04/17/2025).

Viviani, Marco and Gabriella Pasi (Sept. 2017). "Credibility in Social Media: Opinions, News, and Health Information—a Survey". In: *Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery* 7.5, e1209. ISSN: 1942-4795. DOI: 10.1002/widm.1209. (Visited on 04/17/2025).