Issues in Information Systems
Volume 26, Issue 1, pp. 250-264, 2025

DOI: https://doi.org/10.48009/1_iis_119

Using large language models to investigate the origin of chronic wasting disease

G. Kent Webb, *San Jose State University, g.webb@sjsu.edu*

Abstract

This use case reports on the impressive output, hallucinations, instability, and limitations of three Large Langue Models (LLMs): ChatGPT, Gemini, and Grok. The LLMs were prompted in an investigative sequence and responses checked. The collected information supports the common theory that chronic wasting disease (CWD) in North America originated in 1967 at a research facility in Fort Collins, Colorado, where deer were reported to have been exposed to sheep with a similar disease - scrapie. Findings include that: no sheep with scrapie were detected in the area around Fort Collins prior to 1967; domestic sheep reportedly exposed to scrapie were in the facility; there were medical experiments; Fort Collins was active in the scrapie eradication program; three early infection sites, all linked to Fort Collins, are missing from USGS maps showing the disease history; the recently discovered European CWD cases can be explained by local conditions; scrapie in sheep and deer with CWD symptoms were reported centuries ago in Europe; early models simulating disease history lacked adequate data and detail, and ignored the presence of infected captive herds. The LLMs provided good insight into disease simulation, created simulation models, and generated python code.

Keywords: chronic wasting disease etiology, Fort Collins theory, disease simulation, large language model performance, artificial intelligence use case

Introduction

When prompted with the question "What is chronic wasting disease and why is it important?" Gemini, a large language model (LLM), responded with referenced statements including: "Chronic Wasting Disease (CWD) is a fatal neurological disease that affects cervids, which are members of the deer family… It's caused by a misfolded protein called a prion… Potential Risk to Other Species (including humans)." Grok's response included "spreads easily through direct contact, bodily fluids, or contaminated environments ... rapidly evolving … threatens cervid populations." ChatGPT responded, "Prions are extremely resilient, persisting in soil and on surfaces for years … an incubation period of 16 months to 3+ years …" These are representative statements of many sources on the internet referenced by the LLMs so the responses from the LLMs used in this study – ChatGPT, Gemini, and Grok -- were similar and consistent, but not identical when prompts were repeated. Links to source material made them easily accessible for review.

Analysis of prion (misfolded protein) structures has been one of the successful applications of artificial intelligence (AI) (Martinez et al., 2024). The 2024 Nobel prize in Chemistry was awarded for the use of

AI in protein structure prediction. Advances in this area will provide more depth of understanding of the general question addressed here on the origin of CWD. It is a complex topic. Sources on the internet are sometimes contradictory, which seemed to be a cause of instability in some of the LLM responses described in the next sections as they sampled different sources. Zhang et al. (2024) proposed a probabilistic approach that might allow models to infer missing data and resolve inconsistencies.

The leading theory for the origin of the disease has been as described in a New York Times article (Zimmer, 2017): "Scientists long suspected that deer and related species developed chronic wasting disease by picking up scrapie from sheep flocks kept at Colorado State University [Fort Collins]." Also, from the Wall Street Journal (McWhirter, 2019), "Early on, every case of CWD could be traced back to ... the Front Range experimental station in Fort Collins. Infected subjects were unwittingly sent to other states and then released into wild populations." Webb (2021) found published evidence supporting the trace back to Fort Collins as part of a long-term, internet text mining project, also finding that the Fort Collins theory was the most mentioned theory although more sources reported that the origin was unknown. This article with supporting documentation is available with open access on ResearchGate.net.

The recent discovery of CWD in Europe has renewed interest in a spontaneous origin theory. However, there are two types of these prion diseases: classic which is infectious, as with the North American epidemic, and atypical which arises sporadically in older animals and transmits poorly under natural conditions (Greenlee, 2019). Most of the European cases have been atypical. An outbreak among reindeer in Norway is thought to be the result of "antler cannibalism" where reindeer gnawed on and consumed the antlers of other reindeer with the atypical form of the disease (Mysterud et al, 2020). Scrapie in sheep was reported by 1750 or earlier in Europe and a similar disease in deer was reported in Europe by 1772 (Ness et al., 2023) although the authors find that testing and symptoms for some of these reports about deer indicate rabies. Other origin theories such as environmental pollution or bioweapons research have garnered little support and were generally ignored in the analysis by the LLMs.

Methodology

Three LLM's were prompted with a sequence of queries central to this research topic, sometimes suggested by the LLM responses. Table 1 summarizes the sequence of prompts and responses from the LLMs that are discussed in the following sections. Prompts were identical for each LLM and for tests of consistency. LLMs did not respond identically to identical prompts. Each LLM was prompted multiple times during the data gathering period from March to May 2025. Salinas and Morstatter (2024) found that small changes in prompts, even adding a space, could result in a different response. The text for each major prompt appears below as the title of a section.

At Chatgpt.com, GPT-4o was available. The site warned that "ChatGPT can make mistakes. Check important info." Checking proved to be a useful method of discovering additional relevant information during this project. On the Google Gemini website, version 2.0 Flash was available. The site also warned "Gemini can make mistakes, so double-check it." Grok-3 was available on the X.com website and the X phone app. Grok provided the most entertaining narrative of its "thought process" and the most complex mathematical responses.

Results

Table 1 summarizes the sequence of major prompts and responses from the LLMs that are discussed in the following sections. The following section titles are prompts made to the LLMs.

Why Is It Important to Know the Origin of a Disease?

For this prompt the LLMs produced similar responses to those listed in Table 1. Many sources with consistent information were available on the internet and were referenced in the responses. The LLMs could rewrite many combinations of these sources and produce a consistent response   Gemini summarized a list of bulleted items in its response with: "In essence, knowing the origin of a disease empowers us to take proactive steps to protect ourselves and our communities." Grok concluded, "knowing a disease's origin is foundational to controlling its spread, understanding its nature, and preventing future outbreaks." When the prompt was changed from "a disease" to "chronic wasting disease," answers were similar but with added specifics about CWD such that it affects deer and that it is a prion disease. The ChatGPT response included these bulleted headings: "Understanding Transmission Pathways, Improving Control and Prevention, Safeguarding Other Species, Developing Better Surveillance …"

Table 1. Comparison of LLM Responses

Prompt	ChatGPT	Gemini	Grok
Why Is It Important to Know the Origin of a Disease?	"crucial for … Prevention and Control … Avoiding Misinformation"	"empowers us to take proactive steps to protect ourselves and our communities"	"foundational to controlling its spread … and preventing future outbreaks"
What Is the Origin of Chronic Wasting Disease? (Different responses from multiple prompts)	"believed to have emerged in captive mule deer in Colorado" or "a theory" or "leading theory "or "unknown" or "spontaneous a theory"	"unknown" or "Fort Collins leading theory" or "Fort Collins is a theory" or "possibly spontaneous or scrapie" or "scrapie"	"a research facility in Fort Collins" or "derived from scrapie … Fort Collins facility" or "unknown" or "spontaneous"
What disease is the most likely source of CWD?	"most likely linked to scrapie" - a paragraph summary	"the hypothesis of a scrapie origin is compelling" with details	"Scrapie" - a final answer with explanation
When Was Scrapie First Reported in Colorado? (Kiowa County, 1966)	"Colorado in 1947" right year for U.S. but not for Colorado, later "not available"	"unknown" gave the answer (at left) days after the information was posted to the web	"1967 in Weld County" wrong location and year is for CWD, later "speculates" the "1940s or 1950s"
Write a Simulation Model for CWD Disease Spread and Python SciPy?	SIR (Susceptible Infected Recovered) simple model, python code	"a basic SIR model" created Python code, many difficult to quantify variables, python code	SIR with spatial spread on a 2D grid, most complex model, more difficult variables, python code
What is the Basic Reproductive Number for CWD?	Typical 1.2 to 2.0, some studies 1.1 to 3. As high as 10 to 11.	1.3 or greater, one study 2.63 to 4.458 in the early stages for mule deer	one study 2.2 to 4.5 in the early stages for mule deer in Alberta
Analyze Simulation Research Articles	Most concise	Wrote submission ready research. Most polite.	Most detail
Run Simulations with published model or with similar data	Less stable. One unexplained wild result.	Declined sometimes to run the simulation with limited data, used logic?	Most stable. Finding unusual data pattern, used logic? Most complex models

What Is the Origin of Chronic Wasting Disease (CWD)?

At the start of the LLM prompts for this question, responses appeared to be consistent. ChatGPT concluded that the disease is believed to have originated at the Fort Collins facility. Gemini concluded that Fort Collins was the leading theory. Grok declared, "I can say: 'The origin of chronic wasting disease is in Colorado, United States, where it was first observed in captive mule deer in 1967.'" Using a text mining approach, Webb (2021) found some reliable sources not found by the LLMs were newspaper articles, perhaps because they were behind a paywall or no longer available on the internet. Several newspaper articles reported interviews with individuals who had worked at the Fort Collins facility and remembered deer sharing pens with sheep. One researcher there at the time, Gene Schoonveld, reported that the sheep mixed with deer were from a Colorado State University at Fort Collins scrapie research project (Blumhardt, 2018), a key aspect of this analysis.

As research progressed, prompts were resubmitted to the LLMs as a check of consistency. Surprisingly, the responses changed. For example, a ChatGPT response: "Current research points to a combination of spontaneous prion formation and conditions in captive settings as contributing factors …" Gemini responded that the Fort Funston theory "remains unproven" and "the most prominent theories involve either a spontaneous misfolding of the prion protein within cervids or a cross-species transmission event from scrapie-infected sheep." However, in a later response, Gemini again stated that Fort Funston was the leading theory and responses continued to vary over time reflecting conflicting conclusions available on the internet. In another prompt Gemini summarized several websites with this conclusion about how CWD is spread: "the prevailing scientific understanding is that it primarily spreads through transmission rather than arising spontaneously in multiple locations." Grok provided a "Think" with these responses: "The evidence leans toward CWD starting in a research facility" (later identified as Fort Funston), "computer modeling suggests CWD may have been present in free-ranging mule deer since at least the 1960s" (this is the basis for several of the following prompts). In one response Gemini included one of the less often reported theories: "It's also possible that an unknown environmental factor triggered the initial misfolding of prion proteins in deer."

The LLMs occasionally missed important information in the references used for the response and the weighting of the relative likelihood of the theories reflected the sources referenced in the response. Chang et al. (2024) provides a survey of LLM evaluations, noting that LLMs lack abstract reasoning and "are prone to confusion or errors in complex contexts." Based on these limitations, the following more specific prompts were submitted to address the complex overall question of the origin of the disease.

What Disease Is the Most Likely Source of Chronic Wasting Disease?

As reported in Table 1 all three LLMs responded that scrapie, a similar prion disease of sheep, was the most likely disease source. ChatGPT summarized, "This connection is based on studies comparing the biochemical properties and transmission characteristics of the prions involved." Gemini produced a submission ready academic paper titled "The Etiology of Chronic Wasting Disease: A Comprehensive Analysis" with 118 citations including Greenlee et al. (2023) concluding, after inoculating five deer with classic sheep scrapie, that "WTD [White-tailed deer] are susceptible to the agent of classic scrapie from sheep, and differentiation from CWD may be difficult." From Cassman, Frese, and Greenlee (2021), "Our data suggest that the phenotype of CWD in sheep is indistinguishable from some strains of scrapie in sheep. Given our results, current detection techniques would be unlikely to distinguish CWD in sheep from scrapie in sheep." Numerous laboratory studies were cited involving the inoculation of sheep, deer, and other species with CWD or scrapie. Grok made a broader summary, "Based on historical timing, geographical proximity, biological compatibility, and supporting experimental evidence, scrapie is the most likely source of chronic wasting disease."

Gemini commented, “the spontaneous origin hypothesis has gained increasing attention with the discovery of atypical CWD in Europe and the laboratory-based generation of prions.” To clarify, Gemini summarized Greenlee (2019) who explains that there are two types of scrapie, “classical scrapie strains that are readily spread within populations of susceptible sheep... and atypical (Nor-98) scrapie ... is thought to occur spontaneously in older sheep or goats… transmits very poorly under natural conditions” The article observes that classical scrapie “was first described nearly 300 years ago in the United Kingdom and other western European countries.” Gemini noted that CWD in North America spreads similarly to classical scrapie, through direct contact or environmental contamination. When prompted “Does chronic wasting disease spread spontaneously?” Gemini responded, “The atypical cases in Europe might suggest that spontaneous prion misfolding can occur in cervids, but these do not appear to be readily contagious in the same way as classical CWD.” One Gemini response cited Silva (2022) that European prion strains are “different from those in North America.” Grok noted with a reference to Barrio et al. (2024) the cases in Europe, such as Norway “appear to involve different prion strains ... Whereas CWD strains circulating in North America exhibit some uniformity, the cases found in Europe are more variable.”

Mysterud et al. (2020) explain how a transmissible CWD may emerge from atypical CWD by cannibalism, where reindeer in Norway were observed gnawing on and consuming the antlers of other reindeer. This local outbreak of CWD among the reindeer in Norway would be an analog of the Kuru epidemic of a similar human prion disease that apparently spread by ritual cannibalistic practices. Research is rapidly developing in this area.

ChatGPT was prompted with, “When was scrapie first discovered in the United States” and responded with, “in 1947 within a flock of imported British-origin sheep in Michigan” providing several supporting references, adding it was first detected in goats in 1969 in Texas. Gemini gave a similar response adding it “has been recognized in other parts of the world, particularly Europe, for over 250 years.” In checking an LLM reference to Ness et al. (2023, p. 7) it appears that scrapie has been observed in Europe as early as 1693. One writer, Comber, in the late 18-century reported (p. 10): “I am assured by several Persons of Credit, that a Distemper exactly the same as *Rickets* [scrapie] in Sheep is found to have arisen of late Years among Deer.” Testing of some deer in outbreaks turned out to be rabies with other outbreaks having symptoms of rabies, so the authors cavalierly conclude the cases Comber describes are rabies, not CWD. However, in this case, Comber is reporting the observations of several other “Persons of Credit” and large outbreaks would be classic, not atypical.

When Was Scrapie First Reported in Colorado?

None of the LLMs answered this prompt correctly. The information was found on the internet but only after a very labor-intensive, targeted search. As reported in Table 1, ChatGPT and Grok on the first prompt confidently gave incorrect answers; hallucination is the term often used in the literature reviewed by Huang et al. (2025). ChatGPT concluded, “Scrapie was first detected in Colorado in 1947.” Correct date for the U.S., but the location was Michigan. When prompted weeks later it changed its response to “unknown.” On the initial prompt, Gemini concluded the date was unknown, but the response later changed to “scrapie was not first reported in Colorado,” and “first reported date is elusive.” Grok initially got the wrong location and date: "Colorado in 1967 in Weld County.” The year is when CWD was first observed in multiple deer but not yet identified. The county is not correct for CWD or scrapie, but the USGS map, Figure 1a, mistakenly puts Fort Collins in Weld County instead of Larimer County. Perhaps Grok is reading this map. Grok apologized for making a mistake when asked for documentation, then asked for advice. Martino et al. (2023) suggest that including contextual information in the prompt can reduce hallucinations.

Grok responded to the same prompt a week later with “the leading theory is that CWD might have originated from scrapie in sheep … So, if scrapie was present in Colorado before 1967, that could be relevant.” Grok speculated that scrapie would have preceded CWD so would have been in Colorado in the 1940s or 1950s. Grok’s comment that the timing could be relevant inspired a persistent Google search with a variety of keywords finding snippets of text in Google books that were searchable individually but not open to general search. As explained by Sarokin and Schulkin (2016, p. 158) “most of the content is restricted to previews, annoyingly tiny snippet views, or no inside-the-book view at all.”

By combining results from this search, it could be documented that scrapie was first reported in Colorado in Kiowa County (United States Livestock Sanitary Association, 1967, p. 481) in 1966 (New England Sheep and Wool Growers Association,1966, p. 3 and 4; Gill, 2017), about 230 southeast of Fort Collins. This was the only case reported in Colorado in 1965, 1966, or 1967 according to the U.S. Department of Agriculture (1965, 1966, 1967). No scrapie was reported around Fort Collins or southeastern Wyoming before 1964 (Hourrigan, 1964). Wineland (1993) reports that there was scrapie in Larimer County, where Fort Collins is located, by 1991 along with Prowers County and the already mentioned Kiowa County. In the mountain area, including Colorado, by 1991 there were only eleven scrapie infected flocks out of 14,443 for 0.07% of flocks – rare. Wineland was contacted by email to see if the scrapie data used in the article was available, but no response.

From additional Google Books search it was discovered that the chairman of the U.S. Scrapie Eradication Committee was Colorado’s Commissioner of Agriculture, Paul Swisher, an owner of sheep (National Association of State Departments of Agriculture, 1959, p. 14). Other sources not referenced by the LLMS, indicate that Fort Collins was doing scrapie research on sheep in the facility such as Höing and Bennett (2019) and Gerhardt (2001) who reports that several people recall that sheep and deer were mixed in pens and that researcher Gene Schoonveld was told the sheep were from a scrapie project Gerhardt also reported that two researchers who are referenced in the following sections, M.W. Miller, lead author of a simulation to be analyzed, and E. Williams, who first identified the CWD lesions in the brains of deer, years later did not find evidence of scrapie at Fort Collins in the literature or by asking around.

Write A Simulation Model for CWD Disease Spread and Python SciPy?

This and the next sequence of prompts was based on previous LLM responses that reported statements supporting a spontaneous origin theory like this from the National Academies of Sciences, Engineering, and Medicine (2024, p. xiii): “Although first observed as a syndrome in Colorado in 1967 … CWD is likely to have been present for some time prior to this initial observation.” The article references two disease simulations (p. 20) to support that conclusion: Miller et al. (2000) and Wasserberg et al. (2009). Similar statements about simulations in other articles were often made without a source so the LLMs were used to search for this evidence and found another simulation article, Jennelle et al. (2014). As a start to evaluating these three simulation models, the LLMs were prompted to write a model with an added feature of providing Python code to run the model. As Gemini explains “SciPy is a core library for scientific and technical computing in Python.”

As summarized in Table 1, ChatGPT created an “SIR model” which is a basic disease spread model known as Susceptible-Infected-Recovered with the assumption that the disease is spread by contact but suggested including “environmental contamination or other factors influencing transmission.” Gemini also created an SIR model. Grok created a spatial SIR model. All three LLMs generated Python SciPy. The variables in the ChatGPT model were susceptible deer, infected dead deer, total deer, deer death rate, and the transmission rate which is how frequently deer become infected by contact with other deer. Webb (2018) discusses the problems with estimating deer population data.

The estimated value of the transmission rate, also known as the Basic Reproductive Number or R0, is critical for the accuracy of these simulation models. Gemini created a similar SIR model and added that the transmission rate and death rate are important to the outcome and that "you could add environmental contamination, age structure of the deer population, and spatial distribution." Grok created a similar model but added spatial variables representing the movement of deer and a "discrete Laplacian operator, representing spatial diffusion." Grok did give cautionary advice based on Winter and Escobar (2020, p. 741): "Considering that infectious prions may not follow biological rules of well-known wildlife pathogens … assumptions used when modeling other infectious diseases may not apply for CWD."

What Is the Basic Reproductive Number For CWD?

Gemini described R0, the basic reproductive number used as a variable in the simulation models, as "the expected number of new infections that would result from a single infected individual introduced into a completely susceptible population." The number of infections over time is stable when R0 equals one, increases more rapidly as R0 rises above one, and declines more rapidly as R0 falls below one. In exploring this concept, all three LLM summarized numerous articles commenting on the difficulty of estimating this value. Accordingly, it is not surprising that the LLMs found significantly different values in the literature where R0 was estimated from data for CWD infections. ChatGPT responded that the general range of R0 was 1.2 to 3 but cited two articles that emphasized the importance of environmental transmission indicating R0 could be as high as 10 to 11 when considering the role of environmental prion persistence. Gemini found average estimates of 1.3 or greater and cited a study of mule deer in Alberta with an estimate of 2.63 to 4.458. Grok focused on the Alberta study with values of 2.2 to 4.5 which were in the final line in the conclusions section (Potapov, 2015).

A related number used in simulations, often identified as b, is the infectious contacts per infected individual per year or other unit of time. ChatGPT found "1 to 5 infectious contacts per infectious individual per year is a commonly cited baseline in models ... Some more recent models using data from endemic areas suggest values up to 10 or more in high-density populations." Gemini would not provide an estimate explaining that there are so many variables affecting the parameter. Grok agreed with Gemini adding that "A single infected deer can shed ... over 10 mg of infectious prions annually through excreta alone."

Analyze Simulation Research Articles and Run Simulations

The sequence of prompts and responses in this investigation led to three simulations that are supportive of the spontaneous disease theory and were analyzed by the LLMs: Miller et al. (2000) in Colorado (Figure 1a), Wasserberg et al. (2009) and Jennelle et al. (2014) in Wisconsin (Figure 1b and 1c). As noted earlier, the first two simulations were given as evidence for the spontaneous theory in National Academies of Sciences, Engineering, and Medicine (2024). The third simulation was found by the LLMs. Figure 1 shows segments of three maps produced by the United States Geological Survey (USGS) (National Wildlife Health Center, 2025) showing counties with infected free-ranging deer and elk (gray area) and infected captive herds (yellow dots). Red dots and circles represent additional information collected in this study.

Colorado

Figure 1a is a segment from the first map created by USGS showing infections by county and in captive herds prior to the year 2000. Fort Collins was placed a little too far to the southeast on the map and had two infected facilities. The disease was first noticed as an illness in multiple deer in the Fort Collins facilities by 1967 and in elk sharing the same pens in 1979 (Williams and Young, 1982). Fort Collins was exchanging deer and elk with the Kremmling and Meeker facilities in Colorado and the Sybille Research unit (near Wheatland and sometimes referred to as the Wheatland facility) in Wyoming; also, with zoos in Wyoming, where CWD was identified, and in Colorado where symptoms were observed (Willams & Young, 1992, p. 554). Williams and Young also report that all known infected free-ranging deer were found within approximately 50 km of the main Fort Collins facility and two infected elk were found withing

50 km of the Sybille facility. By 1995 all 49 infected deer and elk cases were identified within 100 km of Fort Collins with the first case in the wild in 1981 (Spraker et al., 1997). CWD was diagnosed in a Toronto Zoo in deer transported from the Denver Zoo (Dube, 2006). Figure 1a corresponds to the data used by Miller et al. (2000) which discusses the infected captive herds shown in the Colorado and Wyoming region of Figure 1a but does not include their presence in the simulation model.

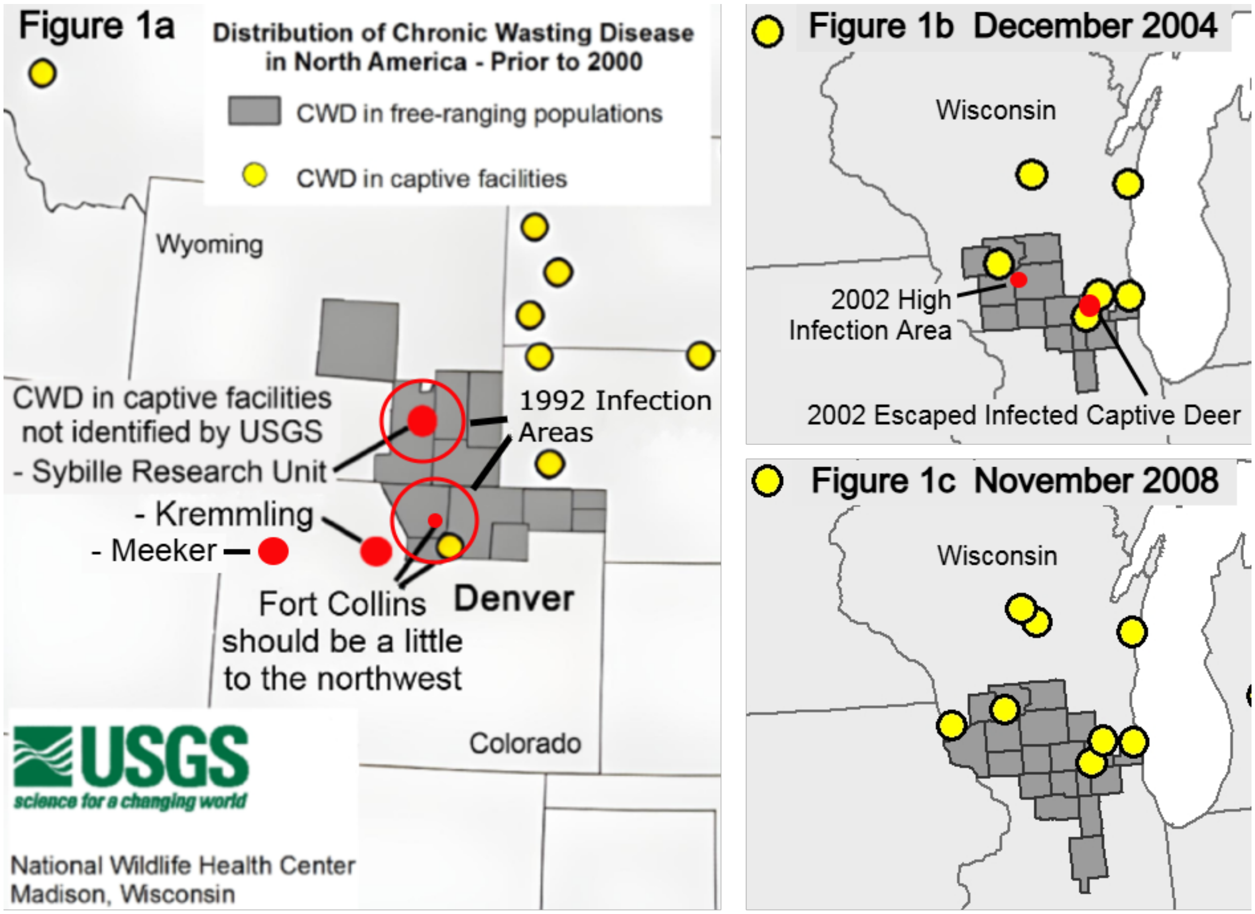

Figure 1. United States Geological Survey (USGS) chronic disease maps (National Wildlife Center, 2025). Discussion and background information is presented here:

The gray in Figure 1 show counties with infected wild herds. Yellow dots represent infected captive herds. Additional information was added to the map in red. The USGS map puts Fort Collins in Weld County, but it is in Larimer County (small red dot). Three infected facilities prior to 2000 not on the USGS map (larger red dots) that received deer and elk from Fort Collins: Sybille (near Wheatland), Kremmling, and Meeker. The first infection discovered in the wild was in 1981. By 1992 the infected areas (red circles) were within 50 km of the main Fort Collins facility for deer and within 50 km of the Sybille facility for two elk. By 1995 all infected wild deer and elk were within 100 km of Fort Collins. Figure 1b shows the first USGS map for Wisconsin after CWD was reported in 2002. Limited testing began in Wisconsin in 1999. Three wild deer taken in the 2001 hunt tested positive in 2002. With expanded testing in 2002, CWD was detected in two captive herds and in an escaped captive deer (red dot, Figure 1b) after several had escaped from an infected Walworth County captive facility. A decayed deer carcass with a captive facility deer tag was found in the 2002 high infection area in Dane county (red dot). Elk was imported to Wisconsin from CWD-affected herds in Colorado during the 1990s and animals were exchanged among local captive facilities. By December 2004, the disease was widespread in southern Wisconsin, Figure 1b, and continued to expand into 2008, Figure 1c. Simulation models investigated here ignored the presence of infected captive herds.

The three LLMs were asked to summarize the model in Miller et al. While the responses were similar, Grok presented the best result based on clarity and detail. The other models agreed with Grok's conclusion that the article "does not explicitly provide detailed simulation equations... The article provides a general equation to model the probability of transition from susceptible (S) to latent (L) states." The model uses a parameter b, discussed in the previous section, defined as the "average number of infectious contacts per infectious individual per unit time." which Grok finds to be a linear relationship to R0 based on a few numeric examples in the article. "From page 686: value of b for 1.1 to 1.3 related to modeled CWD R0 of 1.5 to 1.95." The article about mule deer in Alberta mentioned in the previous section (Potapov, 2015) found values for R0 of 2.2 to 4.5 for mule deer.

Chat GPT summarizes the findings of the simulation: "Epidemics reaching ~15% prevalence in 37–50 years occur at b =1.2 − 1.3." This result would put infected deer in the wild before the infection was observed in the captive facility, suggesting the origin was spontaneous. The data in Table 1 of Miller et al. for 1996 to 1999 show a prevalence (infection rate) for mule deer of 14.7% in the area just north of Fort Collins. Values of b = 1.4 or higher were subjectively discarded because the simulation results did not match the field data used in the study. Grok and ChatGPT were prompted with "Explain the variable values in this simulation model and estimate the years until 15% infection rate is achieved when the R0 value is 3.3", near the middle of the range reported by Potapov (2015). Grok estimated 15 to 18 years, noting the values are approximate since the full model was not presented in the paper. ChatGPT gave a rough estimate of 15 to 20 years but added that increasing R0 "often shortens epidemic timelines considerably" so speculated that 12 to 18 years is "probably" a better answer. These simulations put CWD in the wild well after the captive infections. Gemini concluded the article didn't provide this estimate. Any spillover from the captive herd also reduces the time since the first infection in the wild herd.

Grok concludes that the model "does not account for spatial structure or environmental transmission, which the article notes as potentially relevant." Gemini agrees, specifically adding that deer density is not considered and that it uses a hypothetical deer population of 1,000. ChatGPT responds that the model "does not include human management interventions, social behavior and group structure, and that it is deterministic while random events could drastically alter outcomes in small populations." The small sample size problem is acknowledged in the paper. The field data includes 146 infected wild deer or elk out of 4,108 used in the analysis (Table 1) collected from October 1996 to January 1999 spread over north central Colorado and southeast Wyoming as shown in Figure 1a. Miller et al. discarded earlier infected deer data. As reported earlier, this data is consistent with an outward spread from Fort Collins and Sybille (near Wheatland) as illustrated in Figure 1a. All infected deer were found within 50 km of Fort Collins and two infected elk were found within 50 km of Sybille (Williams and Young, 1992). All infected deer and elk were found within 100 km of Fort Collins by 1995 (Spraker et al., 1997). Miller was second author on the Spraker paper and G.G. Schoonveld, who has reported on the mixing of deer in pens with sheep from a scrapie project, was also a listed author. Miller et al. (2000, p. 687) comments that the highest infection rate should be found near the origin. The highest infection rate in the field data used by Miller et al. was 14.7% for mule deer in the area just north of Fort Collins. Denver is to the southeast, blocking deer movement. The areas south and north of the infected Sybille research facility in Wyoming (Figure 1a) were next highest at 13.1% and 12.4% for mule deer.

As an example of a condition with effectively no variable distance, as in the simulation of Miller et al., Williams and Young (1992, p. 556) report that during the peak of CWD at Fort Collins from 1970 to 1981, 90% (60 deer) died or were euthanized because of the disease (all eventually died, the disease is 100% fatal). Grok was prompted to find the b and R0 values required to reach a 90% infection rate in 15 years, assuming the epidemic started slightly before numerous deer were observed with symptoms in 1967 and running to 1981. Grok noted this "requires intense transmission, more akin to captive settings" and that b would be 6.2 infectious contacts per infectious individual per year with R0 of approximately 8.9.

Another test of the model's predictive power comes from this statement in Miller et al. Figure 5: "Models forecast population declines ... once prevalence exceeded about 5%." When prompted to find research estimating at what prevalence populations decline, Grok found that population declines when prevalence reached around 27% (Edmunds et al., 2016) for whitetail deer. Gemini created a full academic style paper concluding population declines for mule deer in Colorado can begin at around 20% with a higher prevalence required for whitetail deer. ChatGPT agreed with the 20% threshold for mule deer.

Wisconsin

For Wisconsin, Figure 1b is a segment from the second USGS map produced in December 2004, and 1c is a segment from November 2008, corresponding to the data used in Wassenberg et al. (2009). Limited testing began in Wisconsin in 1999 with 1,204 animals tested from 1999 to 2001. Three wild deer taken in the 2001 hunt tested positive in 2002. Also in 2002, CWD was detected in two captive herds and in an escaped captive deer. A tree fell on the fence of an infected, unlicensed facility in Walworth County (N9640 Eagle Rd.) allowing an unknown number of deer to escape (Ahern, 2002). Joly et al. (2003) reports that one of the recaptured deer tested positive for CWD and that elk were imported to Wisconsin from CWD-affected herds in Colorado during the 1990s. Infections in Walworth County (Figure 1b) were apparently reported after Joly et al. (2003). They analyzed a high infection area in the 2002 data in the Mount Horeb area of Dane Country where the Department of Natural Resources (DNR) discovered a decayed deer carcass with a captive facility ear tag (Nohl, 2008). Nohl reports that more than 100 deer were known to have escaped, more than 3,000 animals had been imported to the state with some from infected areas, and 900 animals had been moved among captive facilities within the state.

Wasserberg et al. (2009) presents two simulation models but make no mention of the infected captive herds in Wisconsin (Figure 1b and 1c) or the escape of captive deer. Gemini responded to a prompt to compare Miller et al. with Wasserberg et al. which it described as "a more complex, demographically detailed matrix model, explicitly comparing DD (density dependent) and FD (frequency dependent) transmission." Among the limitations in the simulations, Grok noted "they did not consider the effects of immigration or emigration ... The full integration of spatial dynamics and the complexities of environmental transmission in wild populations continues to be a challenge." Grok suggested that combing the two transmission modes, DD and FD, into a single model could "better reflect real-world dynamics" but that the available data may not be adequate to combine the models." The significantly different outcomes from the two models, DD and FD, reflect on the low reliability of the simulations with time since disease introduction (TDI) at 36 years for DD and 188 years for FD. These results would put CWD origination prior to CWD detection in captive herds, suggesting a spontaneous origin. While ignoring the impact of infected captive herds around the study area (Figure 1b and 1c), Wasserberg et al. (p. 464) concludes that "a larger number of initial infected deer" would shorten the time estimate since the first infection. That would include a spillover from infected captive deer.

Wassenberg et al. had a limited sample size of four years of annual data for a total sample of 7,582 deer and 375 infections for an infection rate of 4.9%. Wasserberg was co-author with Jennelle et al. (2014) of a similar model with a few more data points that found a shorter time since disease introduction at around 40 years, or 1970 based on article discussion. The Jennelle et al. article specifically asserts there has been "no current evidence to suggest that CWD spread in our study area was facilitated by humans" (p. 6). As noted previously for Figure 1b and 1c, there were infected captive herds and escaped captive deer in southern Wisconsin. Joly et al. (2003, p. 600) find "the most commonly suggested hypothesis is that CWD in

Wisconsin may have emerged through importing of an affected cervid." One "hunter" disposing part of an infected deer carcass from a captive herd would be enough to start the spread.

In comparison to Wassenberg et al., Grok responds that Jennelle et al. "Uses data from 2002–2013, providing a longer time series (11 years vs. 4 years)... Explicitly models sex-specific FD transmission... Estimates CWD spread at 1.13 km/year ... density-dependent harvest rates ... and a sophisticated sensitivity analysis." Actually, Jennelle (p. 2) uses "October and January 2002–2011 for parameter estimation ... October and January 2011–2013 for validation." Gemini found other estimates of CWD spread "roughly 6.9635 to 7.3201 kilometers per year" and Grok at "5 to 15 miles per year" with both noting high uncertainty. While the model ignores the impact of infected captive herds, Jennelle et al. state (p. 3) "increasing the initial number of infected deer results in lower time since disease introduction (TDI) estimates." The model estimates different, but similar TDIs for different regions in the area.

The LLMs were used to run numerous simulations using infection sample data for the southern farmland zone and population estimates. Using just two data points of CWD positives divided by sample size for 2002 and 2010 (Wisconsin Department of Natural Resources, n.d.) in the endemic southern farmland zone and a simple logistic or exponential growth model, ChatGPT and Grok produced estimates like Jennelle et al. of 1970 and 1971, respectively. Using annual CWD prevalence from 2001 to 2010 and a population estimate of 27,500 (25,000 to 30,000 for the area from Nohl, 2008), Grok estimated the origin year of 1995 with the disease growing exponentially. Gemini estimated 1998. ChatGPT used a logistic growth curve to estimate 1981. Somewhat different answers were reported as prompts were repeated because LLMs tried different growth models. In one case ChatGPT estimated 998 A.D. without explanation. There were some responses that seemed to represent judgment. Several times Gemini declined to make a modeled estimate because of the small sample size and deduced the origin to be in the late 1990s given the small number of infections in 2001. Grok warned that an unusual increase in infections beginning around 2008 indicated something else was influencing the process.

Gemini responded to the prompt for a comparison of Wasserberg et al. to Jennelle et al. that the later had "sex-specific infection rates, stronger empirical support for frequency-dependent transmission, more biologically plausible and refined estimates of TDI, and the initial quantification of geographic disease spread." (1.13 km per year, Jennelle et al., p. 5). This low transmission rate does not account for the wide geographical distribution of infections in southern Wisconsin, presumably attributed to multiple spontaneous outbreaks. From the National Academies of Sciences, Engineering, and Medicine (2024, p. 64) referring to Wisconsin: "Field data illustrate spatial expansion of what likely are multiple epicenters over time." The statement referred to Figure 5.2 showing cumulative infected wild animals for years 2002, 2012, 2022 concentrated in southern Wisconsin. Multiple spontaneous outbreaks over a relatively narrow time and geography seem unlikely. Walter et al. (2025) found the CWD prions in the Wisconsin area to be similar, indicating a common origin. While this issue was not discovered by the LLMs, they were prompted with "How likely is it that chronic disease spontaneously originated in multiple locations in Wisconsin?" Grok responded, "highly unlikely ... without a clear causal link." Gemini agreed commenting that it was a "fascinating question." ChatGPT agreed commenting it was "highly unlikely" without a common cause. M.W. Miller, who was lead author on the Colorado simulation, served on the committee for this National Academies of Sciences, Engineering, and Medicine publication.

Discussion

The LLMs provided tantalizing research support: fast, elegant, and occasionally sloppy. Although the error rate was small given the amount of response generated for this project, it required careful checking. However, checking led to the discovery of important information. Mistakes or variations in response were typically related to missing or contradictory information. The first discovery of scrapie in Colorado – reported as Kiowa County, 1966, about 230 miles southeast of Fort Collins – was available but hard to access on the internet. For the fundamental question of disease origin, there has been a new push for the spontaneous theory resulting in contradictory available information. Historically this theory seemed to reappear each time CWD was found in a new location as in Wisconsin. Recently, problematic simulations that ignored the presence of CWD in captive herds and discovery of CWD in Europe with explainable origins, were used to revive the spontaneous creation theory for the North American epidemic. The theory provides deniability for the culpability of the captive deer complex and governing organizations (for a partial list see National Academies of Sciences, Engineering, and Medicine, 2024).

The first prompt in this investigation was "Why Is It Important to Know the Origin of a Disease?" with similar responses including from Grok: "knowing a disease's origin is foundational to controlling its spread." Many state wildlife agencies have imposed safety measures on captive herds to control CWD, but these restrictions are costly and often resisted by the captive deer industry and their associates with the support of state legislators as in Wisconsin (Nohl, 2008). This paper honors the efforts of Gene Schoonveld who was a deer researcher in the 1960s at Fort Collins and frequently reported (as in Blumhardt, 2018) on the mixing of deer in pens with sheep from a scrapie project, consistent with Colorado's leadership in the scrapie eradication program.

References

Ahern, S. (2002, October 25). Deer hunters unable to locate missing deer so far. *Waukesha Freeman*.

Barrio, T., Benestad, S.L., Douet, J., Huor, A., Lugan, S., Aron, N. & Andreoletti, O. (2024). Zoonotic Potential of chronic wasting disease after adaptation in intermediate species. Emerging Infectious Diseases, 30(12), 2691-2694. https://doi.org/10.3201/eid3012.240536

Blumhardt, M. (August 28, 2018). Deadly Deer Disease. *Windsor Beacon.* W-2. Newspapers.com

Cassmann, E. D., Frese, R. D., & Greenlee, J. J. (2021). Second passage of chronic wasting disease of mule deer to sheep by intracranial inoculation compared to classical scrapie. *Journal of Veterinary Diagnostic Investigation*, official publication of the American Association of Veterinary Laboratory Diagnosticians, Inc, 33(4), 711–720. https://doi.org/10.1177/10406387211017615

Chang, Y. et al. (2024). A survey on evaluation of large language models. *ACM Transactions on Intelligent Systems and Technology*, 15(3). Article No.: 39, 1 - 45. https://doi.org/10.1145/3641289

Dubé, C., Mehren, K.G., Barker, I.K., Peart, B.L. & Balachandran, A. (2006). Retrospective investigation of chronic wasting disease of cervids at the Toronto Zoo, 1973-2003. *Canadian Veterinary Journal,* Dec;47(12):1185-93. PMID: 17217088

Edmunds, D.R., Kauffman, M.J., Schumaker, B.A., Lindzey, F.G., Cook, W.E., Kreeger, T.J., Grogan, R.G. & Cornish, T.E. (2016). Chronic wasting disease drives population decline of white-tailed deer. *PLoS ONE*, 11(8): e0161127. https://doi.org/10.1371/journal.pone.0161127

Gerhardt, G. (2001, November 5). Possible origins of chronic wasting disease. *Rocky Mountain News*, Colorado, USA.

Gill, C. (2017, October 17). Chronic wasting disease (CWD) - America's answer to mad cow disease. *Animal Development.* https://www.pitchstonewaters.com/chronic-wasting-disease-cwd-americas-answer-to-mad-cow-disease/

Greenlee, J.J. (2018). Review: Update on classical and atypical scrapie in sheep and goats. *Veterinary Pathology*, 56(1):6-16. doi:10.1177/0300985818794247

Greenlee, J.J., Moore, S.J., Cassmann, E.D., Lambert, Z.J., Kokemuller, R.D., Smith, J.D., Kunkle, R.A., Kong, Q. & Greenlee, M.H.W. (2023). Characterization of classical sheep scrapie in white-tailed deer after experimental oronasal exposure. *The Journal of Infectious Diseases*, 227(12), 15 June 1386–1395.

Höing, A. & Bennett, A.M. (2019). *Humans and Animals: Intersecting Lives and Worlds.* Netherlands: Brill.

Hourrigan, J.L. (1964). "Scrapie in the United States and the scrapie eradication program. *Report of Scrapie Seminar Held in Washington D.C.* January 27-30. USDA Agricultural Research Service

Huang, L. (2025). A survey on hallucination in large language models: Principles, taxonomy, challenges, and open questions. *ACM Transactions on Information Systems*, 43(2). Article No. 42. https://doi.org/10.1145/3703155

Jennelle, C.S., Henaux, V., Wasserberg, G., Thiagarajan, B., Rolley, R.E. & Samuel, M.D. (2014). Transmission of chronic wasting disease in Wisconsin white-tailed deer: implications for disease spread and management. *PLoS One,* Mar 21;9(3): e91043. doi: 10.1371/journal.pone.0091043.

Joly, D.O., Ribic, C.A., Langenberg, J.A., Beheler, K., Batha, C.A., Dhuey, B.J., Rolley, R.E., Bartelt, G., Van Deelen T.R. & Samuel, M.D. (2003). Chronic wasting disease in free-ranging Wisconsin White-tailed Deer. *Emerg Infectious Diseases,* May;9(5):599-601. doi: 10.3201/eid0905.020721.

Martinez, C., Teijeira, S., Domínguez, P., Campanioni, S., Busto, L., González-Nóvoa, J. A., Alonso, J., Poveda, E., San Millán, B. & Veiga, C. (2024). Correlating histopathological microscopic images of creutzfeldt–jakob disease with clinical typology using graph theory and artificial intelligence. *Machine Learning and Knowledge Extraction*, 6(3), 2018-2032. https://doi.org/10.3390/make6030099

Martino, A., Iannelli, M. & Truong, C. (2023). Knowledge injection to counter large language model (LLM) hallucination. In: Pesquita, C., et al. The Semantic Web: ESWC 2023 Satellite Events. *Lecture Notes in Computer Science*, vol 13998. Springer, Cham. https://doi.org/10.1007/978-3-031-43458-7_34

McWhirter, C. (2019, February 7). Disease threatening deer population has spread to 26 states. *Wall Street Journal*, WSJ.com.

Miller, M.W., Williams, E.S., McCarty, C.W., Spraker, T.R., Kreeger, T.J., Larsen, C.T. & Thorne, E.T. (2000). Epizootiology of chronic wasting disease in free-ranging cervids in Colorado and Wyoming. *Journal of Wildlife Diseases*. Oct;36(4):676-90. doi: 10.7589/0090-3558-36.4.676. PMID: 11085429.

Mysterud, A, Ytrehus, B., Tranulis, M.A. et al. Antler cannibalism in reindeer. *Scientific Reports.* 10, 22168 (20200). https://doi.org/10.1038/s41598-020-79050-2

National Academies of Sciences, Engineering, and Medicine. (2024). *State of knowledge regarding transmission, spread, and management of chronic wasting disease in U.S. captive and free-ranging cervid populations*. Washington, DC: The National Academies Press. https://doi.org/10.17226/27449 .

National Association of State Departments of Agriculture. *Annual Convention.* 1959

National Wildlife Health Center (2025). Expanding Distribution of Chronic Wasting Disease. *United State Geological Service (USGS)*. https://www.usgs.gov/centers/nwhc/science/expanding-distribution-chronic-wasting-disease

Ness, A., Aiken, J., & McKenzie, D. (2023). Sheep scrapie and deer rabies in England prior to 1800, *Prion, 17:1*, 7-15. DOI: 10.1080/19336896.2023.2166749

New England Sheep and Wool Growers Association (1966). *Volume 11.* United States: A. A. Lund Associates

Nohl, M.V.K. (2008, January 31). The Killer Among Us. *Milwaukee Magazine*. https://www.milwaukeemag.com/TheKillerAmongUs/

Otero, A., Velasquez C.D., McKenzie, D. & Aiken, J. (2023). Emergence of CWD strains. *Cell and Tissue Research*, Apr;392(1):135-148. doi: 10.1007/s00441-022-03688-9.

Potapov, A., Merrill, E., Pybus, M., & Lewis, M.A. (2015). Empirical estimation of R0 for unknown transmission functions: The case of chronic wasting disease in Alberta. *PlosOne*, October 9. https://doi.org/10.1371/journal.pone.0140024

Salinas, A. & Morstatter, F. (2024). The butterfly effect of altering prompts: How small changes and jailbreaks affect large language model performance. *arXiv preprint,* arXiv:2401.03729.

Sarokin, D. & Schulkin, J. (2016). *Missed information: Better information for building a wealthier, more sustainable future*. United Kingdom: MIT Press.

Silva, J.S. (2022). Chronic wasting disease (CWD) in cervids and the consequences of a mutable protein conformation. *ACS Omega*, 7(15) 12474–12492 https://pubs.acs.org/doi/10.1021/acsomega.2c00155

Spraker T.R., Miller, M.W., Williams, E.S., Getzy, D.M., Adrian, W.J., Schoonveld, G.G., Spowart, R.A., O'Rourke, K.I., Miller, J.M. & Merz, P.A. (1997). Spongiform encephalopathy in free-ranging mule deer (Odocoileus hemionus), white-tailed deer (Odocoileus virginianus) and Rocky Mountain elk (Cervus elaphus nelsoni) in northcentral Colorado. *Journal of Wildlife Diseases,* 33:1–6. doi.org/10.7589/0090-3558-18.4.465

United States Livestock Sanitary Association (1967). *Proceedings, Annual Meeting of the United States Live Stock Sanitary Association.*

U.S. Department of Agriculture: (1965, 1966, 1967). Animal Morbidity Report. *Veterinary Services, Animal & Plant Health Inspection Service.* https://www.google.com/books/edition/Animal_Morbidity_Report/6Av_wY6ZzJYC?hl=en&gbpv=0

Walter, W. D., Herbst, A., Lue, C.-H., Bartz, J. C., & Hopkins, M. C. (2025). Overview of North American isolates of chronic wasting disease used for strain research. *Pathogens*, 14(3), 250. https://doi.org/10.3390/pathogens14030250

Wasserberg, G., Osnas, E.E., Rolley, R.E. & Samuel, M.D. (2009). Host culling as an adaptive management tool for chronic wasting disease in white-tailed deer: a modelling study. *Journal of Applied Ecology*, 46: 457-466. https://doi.org/10.1111/j.1365-2664.2008.01576.x

Webb, G.K. (2018). Searching the internet to estimate deer population trends in the U.S., California, and Connecticut. *Issues in Information Systems*, 19(2):163- 73. https://doi.org/10.48009/2_iis_2018_163-173

Webb, G.K. (2021). Text mining to identify the origin of chronic wasting disease. *Issues in Information Systems,* (22)3, 184-196. https://doi.org/10.48009/3_iis_2021_200-213. Documentation available with this article on Researchgate.net

Williams, E.S. & Young, S. (1980). Chronic wasting disease of captive mule deer: A spongiform encephalopathy. *Journal of Wildlife Diseases,* 16: 89–98.

Williams, E.S. & Young, S. (1982). Spongiform encephalopathy of rocky mountain elk. *Journal of Wildlife diseases,* 18(4): 465–471. https://doi.org/10.7589/0090-3558-18.4.465

Williams, E.S. & Young, S. (1992). Spongiform encephalopathies of Cervidae. *Scientific and Technical Review Office of International Epizootics,* 11*:* 551 - 567. DOI: 10.20506/rst.11.2.611

Wineland, N.E., (1993). Epidemiology of reported scrapie in the United States: 1947-1991. Master's Thesis, Clinical Sciences Department. Spring. *Colorado State University, Fort Collins, Colorado*. https://mountainscholar.org/items/ced70898-333c-4de0-84dd-42ff4fc751ef

Winter, S.N. & Escobar, L.E. (2020). Chronic wasting disease modeling: An overview. *Journal of Wildlife Diseases*, 56 (4): 741–758. doi: https://doi.org/10.7589/2019-08-213

Wisconsin Department of Natural Resources (n.d.). *Summary of CWD Statewide Surveillance*. https://apps.dnr.wi.gov/cwd/summary/zone

Zhang, L., et al. (2024). Grounding large language models in real-world environments using imperfect world models. *Authorea*, October 07. DOI: 10.22541/au.172832793.37033556/v1

Zimmer, C. (2017, June 26), Fire may be the only remedy for a plague killing deer and elk. *New York Times*. https://www.nytimes.com/2017/06/26/science/chronic-wasting-disease-deer-elk-prions.html