# Title:

**Reporting and characterization of assumptions required for causal inference in case-control studies on health effects of animal feeding operations: A scoping review protocol**

**What do case-control studies investigating health outcomes among people living near animal feeding operations report and estimate?A scoping review**

# Registration:

This protocol will be made available online at Systematic Reviews for Animals and Food (SYREAF) (www.syreaf.org).

# Authors and Contributions:

**B. Alexander Fonseca Martinez** conceived the idea and developed the protocol.

**Annette M. O’Connor** (oconn445@msu.edu) conceived the idea and developed the protocol.

**Jan M. Sargeant** conceived the idea and developed the protocol.

**Sarah Totton** provided critique and refinement of the protocol.

**Chong Wang** provided critique and refinement of the protocol

# Amendments:

None to report

# Support: source, sponsor, and role of funder:

BAFM was financially supported by the grant #19-146 from the National Pork Board (NPB). Other financial support was provided by AMOC’s discretionary funding. The funders have no role in the design of the study; in the collection, analyses, or interpretation of data; in the writing of the manuscript, or in the decision to publish the results. The other authors declare that they have no conflicts of interest.

# Reporting Guidelines:

No reporting guidelines specific to scoping review protocols are available, therefore, we used a combination of the PRISMA-P 1 and PRISMA-ScR 2 to report this protocol.

# 1. INTRODUCTION

## 1.1. Rationale

Research on this topic of community health and exposure to large animal feeding operations are often observational. Studies in this area focuses on two primary objectives: determining whether there is a higher prevalence of chronic diseases in areas with AFOs, and establishing whether the incidence of diseases is higher in exposed communities compared to non-exposed ones. This work often incudes case-control studies, an epidemiological design characterized by sampling based on the outcome, which can be either incidental or prevalent. It is commonly understood that the odds ratio is the primary measure of effect for case-control studies. 3,4 However, this is specifically true for a particular type of case-control design where controls are sampled from the survivors at the end of the period of interest. The measure of effect reported in case-control studies, typically the cross-product obtained from a two-by-two table or logistic regression, can also be interpreted as the incidence density ratio (IDR), cumulative incidence ratio (CIR), incidence odds ratio (IOR), or prevalence odds ratio (POR) based on the study design depending on many factors. 3–6 To determine if the cross-product obtained from the analysis of case-control studies is mathematically equivalent to the IDR, CIR, IOR, or POR, it is necessary to consider method used for control sampling, the source population giving rise to the study population, and the nature of the cases. Knowing what is estimated is important because impacts how the effect size can be combined with other studies to make inferences about the causal effect. However, such information would best be reported by authors, rather than inferred by reviewers. In our experience, however, such information is often missing from case-control study reports. Therefore, the aim of this scoping review is to assess the reporting of these assumptions crucial for estimating causally significant metrics in case-control studies focusing on the health of communities near AFOs and to report our own characterization of what could have been estimated if the authors did not report the necessary information. The specific review questions are: 1) Which assumptions are discussed in case-control studies investigating the effect of AFOs on the health of nearby residents? and 2) What effect size measures could have been estimated if the authors do not report or discuss the necessary assumptions for estimating causally significant metrics in case-control studies focusing on the health of communities near AFOs?

# 2. METHODS

## 2.1. Eligibility criteria

Table 1 outlines the explicit inclusion and exclusion criteria to be applied in this scoping review. The literature considered will be limited to case-control studies. Studies will be classified as case-control based on the investigators' description of the design, or if not provided, the description in the Materials and Methods sections will be utilized.

Table 1: Scoping review inclusion and exclusion criteria.

|  |  |  |
| --- | --- | --- |
|  | **Inclusion criteria** | **Exclusion criteria** |
| Population | * Humans living in communities near AFOs that might be described as industrial, large, concentrated, or other synonyms.
 | * Production systems that appear to be grass-based, nomadic, or confined smallholder operations based on the authors’ description.
* People who actively participate in livestock production and who are therefore occupationally exposed.
 |
| Exposure | * Any strategy used to measure exposure to AFOs such as odor intensity, levels of contaminants in the air, soil, or water, proximity measured by distance, or AFO animal density units.
 | * Models of AFOs exposure. The relevance of such models of exposure in real life is often unclear.
 |
| Outcome | * Outcomes of interest will be prevalent health state measured on humans. The outcome does not need to be a disease; for example, colonization or culture of bacteria from a human is an eligible outcome.
 | * Self-reported health states are not eligible unless the primary research authors provide evidence of appropriate psychometric properties (validity, reliability, responsiveness) and clinical interpretability (validated).
* Outcomes that do not represent direct health measures in humans (e.g., antimicrobial resistance patterns in soil or water resources).
 |
| Geographical location/country of population: | * All geographic locations are eligible.
 | N/A |
| Publication type | * Case-control studies.
 | * Ecological study designs, descriptive studies (e.g., case reports and series), or other analytic observational studies (e.g., cohort, prevalent studies).
 |
| Timeframe | * 1st October 2014 – 1st October 2023
 | N/A |
| Language | * Any language
 | N/A |

## 2.2. Information Sources

Electronic searches of MEDLINE®(via Web of Science) (2014 – 2023), CABI Global Health (via Web of Science) (2014 –2023), Centre for Agricultural Biosciences (CAB) Abstracts (via Web of Science) (2014 – 2023), and Science Citation Index (via Web of Science) (2014 – 2023) will be conducted.

## 2.3. Search Strategy

Table 2: Search strategy in MEDLINE®.

|  |  |
| --- | --- |
| Search line | Search string |
| 1 | MH=animal husbandry |
| 2 | MH=housing, animal |
| 3 | MH=animal feed |
| 4 | TS=((animal$ OR bovine OR cow OR cows OR cattle OR beef OR pig OR pigs OR piglet\* OR pork OR swine OR porcine OR hog OR hogs OR finisher\* OR sheep OR murine OR lamb OR lambs OR poultry OR chicken\* OR hen OR hens OR broiler\* OR turkey\* OR livestock OR "live stock" OR intensiv\* OR industrial\* OR confined OR confinement OR concentrated OR large-scale) NEAR/3 ("feed\* facilit\*" OR "feed\* operation\*" )))  |
| 5 | TS=(cafo OR cafos OR afo OR afos) |
| 6 | TS=("feed lot$" OR feedlot\* OR feedyard\* OR "feed yard\*") |
| 7 | TS=((animal$ OR bovine OR cow OR cows OR cattle OR beef OR pig OR pigs OR piglet\* OR pork OR swine OR porcine OR hog OR hogs OR finisher\* OR sheep OR murine OR lamb OR lambs OR poultry OR chicken\* OR hen OR hens OR broiler\* OR turkey\* OR livestock OR "live stock") NEAR/0 (operation\* OR facility OR facilities OR confined OR confinement )) |
| 8 | TS=((confined OR confinement) NEAR/2 (feed or feeding)) |
| 9 | TS =((intensive or intensively or large-scale or industrial) NEAR/2 (farm or farms or farming or livestock or "live stock")) |
| 10 | TS=(("animal production" or "livestock production" or "live stock production") NEAR/0 (operation\* OR facility OR facilities)) |
| 11 | #10 OR #9 OR #8 OR #7 OR #6 OR #5 OR #4 OR #3 OR #2 OR #1  |
| 12 | MH= (Environmental Health) |
| 13 | MH= (Environmental Exposure OR Inhalation Exposure) |
| 14 | MeSH HEADING: (environmental pollutants) |
| 15 | MeSH HEADING:exp: (air pollutants) |
| 16 | MH=(water pollutants) |
| 17 | MH=(Environmental Illness) |
| 18 | TS= ("public health\*" OR "environmental health\*" OR "environmental medicine" OR "community health\*")  |
| 19 | SO= ("public health\*" OR "environmental health\*" OR "environmental medicine" OR "community health\*")  |
| 20 | TS= ((community or communities or resident\* or residence$ or neighbor\* or neighbour\* or family or families or local$ or populace$ or school$ or preschool\* or highschool\* or nursery or nurseries or playgroup\* or "play group\*" or kindergarten\*) NEAR/4 (health or disease$ or impact\* or effect$ or exposure$ or expose$ or outcome$ or symptom$ or risk$))  |
| 21 | TS= ((public or community or communities or resident\* or residence$ or living or neighbor\* or neighbour\* or family or families or local$ or population$ or populace or school$ or preschool\* or highschool\* or nursery or nurseries or playgroup\* or "play group\*" or kindergarten\*) NEAR/4 (proximity or vicinity or location$ or located or nearby or "near" or close or closely))  |
| 22 | #21 OR #20 OR #19 OR #18 OR #17 OR #16 OR #15 OR #14 OR #13 OR #12  |
| 23 | #22 AND #11  |
| 24 | MeSH HEADING:exp: (animals) |
| 25 | MeSH HEADING: (humans) |
| 26 | #24 NOT #25 |
| 27 | #23 NOT #26 |

## 2.4. selection of sources of evidence

Study selection has three levels: a first level based on assessing information in titles and abstracts, and a second level based on assessing information from the full text of studies. Screening will be conducted using DistillerSR® (Evidence Partners, Ottawa, ON, Canada). Two independent reviewers will conduct the selection process (ST, BAFM) and disagreements will be resolved by consulting a third expert reviewer (AMOC).

In the first round of study selection screening, the abstracts and titles will be screened for eligibility using the following question:

* Does the title and/or abstract describe an observational study reporting the association between relevant AFOs and measures of health in surrounding-community members?

Each citation that passes level 1 screening will progress to level 2. During this full-text screening, any disagreements will be resolved by consensus between the two reviewers. A third reviewer will arbitrate when consensus cannot achieved (AO). For full-text screening the following questions will be used:

* Is the full text available in English?
* Does the study report a comparative association between a relevant animal feeding operation and measures of health in surrounding-community members?
* Does the study assess the relationship between outcome and exposure at the individual human level?
* Does the study report animal feeding operations that would be reasonably considered either large, concentrated or intensive by modern standards (not nomadic, smallholder or pastoral)?
* Does the study include more than one unit of measurement of exposure?
* Does the study include at least one human health outcome measured using either an eligible survey instrument, test, assay or diseases measure obtained from medical records?

Level 3

* Is the study a case-control study?

## 2.5. Data Charting Process

A data collection form was developed within DistillerSR® to gather relevant data. The form underwent a pretest by two reviewers across the 15 references included until 2014. Subsequently, two reviewers will independently extract the data from all relevant articles utilizing this form. Any discrepancies will be resolved through discussion, and if consensus cannot be reached, a third reviewer will be consulted (AMOC). Information will solely be gathered from the articles themselves; no attempts will be made to contact study investigators for additional or confirmed data. Any missing data will be recorded as 'Not reported', and no assumptions will be made about the unreported information.

## 2.6. Data Items

For each relevant prevalence study identified, two reviewers will extract the year(s) the study was conducted, the study population’s location, the animal species at the AFOs, and a description of the human community (e.g., “neighboring residents of animal farms in the Dutch provinces of Noord-Brabant and Limburg”). The reviewers also extracted each exposure and outcome pair and the effect size reported. The extracted outcomes will be categorized into broader groups: lower respiratory conditions, upper respiratory conditions, antimicrobial resistance, etc.

For each extracted exposure-outcome pair and its effect size, we will first assess whether the authors indicated that the odds ratio (OR) might estimate something other than the OR. We will also evaluate if they discussed the key assumptions required for such an inference, including the control sampling method, the source population, underlying assumptions, and the nature of the cases. Second, we will identify what could have been estimated if the authors did not report and discuss the necessary assumptions. We will follow the decision tree developed by Knol et al. (2008) 5, a method previously utilized by other researchers to infer the effect size measure estimated in case-control studies. 5,7

Firstly, reviewers will determined if the outcome was incident or prevalent . Cases will be classified as prevalent if the outcome does not consider the temporality of the event (e.g., existing allergies), and as incident if the outcome measures a new event within a specific time frame (e.g., wheezing within the past year).

Secondly, reviewers will assess whether the population was likely to be dynamic or fixed. Dynamic populations allow for entry and exit of participants based on defined criteria (e.g., location), while fixed populations not permit changes in membership over time.

For incidence cases in dynamic populations, reviewers will evaluate the next questions:

1. Were controls sampled each time a case occurred?
2. Was the distribution of exposure stable in the source population?

Only if the answer to the first question is no, the second question will be evaluated. If the response to the second question is also no, the cross-product of the exposure-outcome pair will be interpreted as an incidence odds ratio when the disease prevalence is greater than 10%, or as a CIR when the disease prevalence is less than 10%. If the answer to either question is yes, the cross-product of the exposure-outcome pair will be classified as providing an estimate of the IDR.

For incidence cases in fixed cohorts:

Since controls can be selected using three methods — at the beginning of follow-up, at the end of follow-up, or concurrently — the questions evaluated for each method will be, respectively: :

1. Is incomplete information about study participants unrelated to exposure?
2. Is the health event rare (<10% in the exposed group)?
3. Was matching on time considered in the analysis?

If the answer to any of these questions is negative, the effect measure of the exposure-outcome pair will be interpretated simply as an odd ratio. Otherwise, the outcome-exposure pair will be classified as providing an estimate of CIR for question 1 and 2, and IDR for question 3. For prevalent cases (prevalence case-controls studies) the odds ratio always is equal to the prevalence odds ratio, however, if its interpretation will be an IDR if the average duration of disease is the same in the exposed and nonexposed groups and a prevalence ratio (PR) if the disease is rare. 5

## 2.7. Critical appraisal of individual sources of evidence

As this is a scoping review, critical appraisal of the included studies will not be performed.

## 2.8. Synthesis of results

We will use descriptive statistics to summarize frequencies, of case-control studies reporting of critical assumptions.

# BIBLIOGRAPHY

1. Moher D, Shamseer L, Clarke M, et al. Preferred reporting items for systematic review and meta-analysis protocols (PRISMA-P) 2015 statement. *Syst Rev*. 2015;4:1-9.

2. Tricco AC, Lillie E, Zarin W, et al. PRISMA extension for scoping reviews (PRISMA-ScR): checklist and explanation. *Ann Intern Med*. 2018;169(7):467-473.

3. Vandenbroucke JP, Pearce N. Case-control studies: basic concepts. *Int J Epidemiol*. 2012;41(5):1480-1489. doi:10.1093/ije/dys147

4. Labrecque JA, Hunink MMG, Ikram MA, Ikram MK. Do case-control studies always estimate odds ratios? *Am J Epidemiol*. 2021;190(2):318-321.

5. Knol MJ, Vandenbroucke JP, Scott P, Egger M. What do case-control studies estimate? Survey of methods and assumptions in published case-control research. *Am J Epidemiol*. 2008;168(9):1073-1081. doi:10.1093/aje/kwn217

6. Vandenbroucke JP, Pearce N. Incidence rates in dynamic populations. *Int J Epidemiol*. 2012;41(5):1472-1479. doi:10.1093/ije/dys142

7. Cullen JN, Sargeant JM, Makielski KM, O’Connor AM. The case-control design in veterinary sciences: A survey. *Prev Vet Med*. 2016;134:179-187. doi:10.1016/j.prevetmed.2016.10.008