

1 **Title: Characterization of the approach to controlling for biasing**
2 **pathways in epidemiological studies of residential exposure to animal**
3 **feeding operations and human health: A scoping review protocol**
4

5 Registration:

6 This protocol will be made available online at Systematic Reviews for Animals and Food (SYREAF)
7 available at: www.syreaf.org
8

9 Authors and Contributions:

10 B. Alexander Fonseca Martinez conceived the idea and developed the protocol.
11 Annette M. O'Connor (oconn445@msu.edu) conceived the idea and developed the protocol.
12 Jan M. Sargeant conceived the idea and developed the protocol.
13 Sarah Totton provided critique and refinement of the protocol.
14 Chong Wang provided critique and refinement of the protocol.
15

16 Amendments:

17 None to report.
18

19 Support: Source, sponsor, and role of funder:

20 This project is funded by the National Pork Board, Des Moines, IA, USA by the grant #19-146. The
21 funding agency will not influence the design, implementation, analysis of the work, review, or
22 development of the manuscript prior to submission.
23

24 Conduct guidelines and Reporting Guidelines:

25 No reporting guidelines specific to scoping review protocols are available, therefore, we used a
26 combination of the PRISMA-P 1 and PRISMA-ScR2 to report this protocol.
27
28

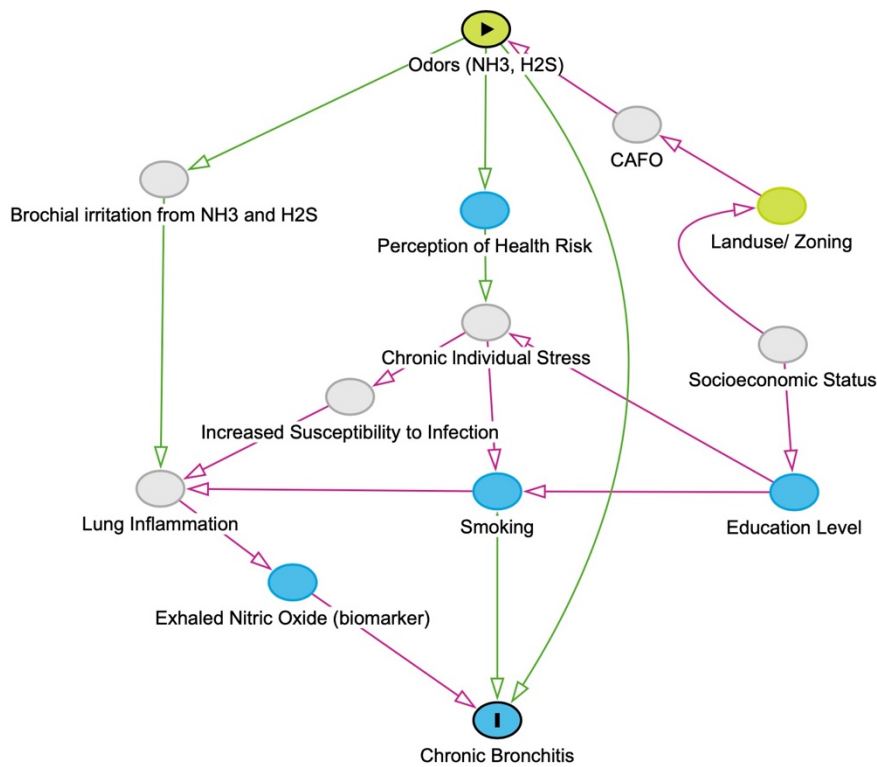
29 **1. Introduction**

30 *1.1. RATIONALE*

31 Animal feeding operations (AFOs) are large industrial-scale farms characterized as an agricultural
32 enterprise where animals are kept and raised in confined situations. This type of operation amasses
33 animals, feed, manure and urine, dead animals, and production operations on a small land area. The
34 increase in the number of AFOs over time has facilitated access to foods of animal origin; however, the
35 adverse effects of large-scale production on the environment and public health remain controversial and
36 debated. Researchers have used observational studies to evaluate the association between living in
37 proximity to an AFO and community health. However, some authors suggest that heterogeneity of
38 outcome definitions and multiple sources of biases prevent reaching a conclusion about the causality of
39 residential exposure to AFOs and adverse health outcomes.

40 While the total effect of exposure refers to the overall impact of a causal factor on an outcome (It
41 includes both direct and indirect effects.), indirect effects refer to the impact of a factor on an outcome

42 that occurs through its influence on other intermediate variables. These concepts impact the adjustment
 43 set of variables that should be controlled in an observational study. Overadjustment bias is induced in
 44 the estimation of the effect by adjusting for an intermediate variable or a descendent of the
 45 unmeasured intermediate variable. These concepts and potential biases are routinely assessed in
 46 directed acyclic graphs (DAGs) which identify the variables that must be adjusted in order to estimate
 47 either the direct or total causal effect. Figure 1 reports a DAG created by the EPA to assess the impact of
 48 AFOs on lower respiratory disease in people living in the surrounding area.



49 Figure 1. Directed Acyclic Graph (DAG) proposed by3. The DAG was generated using DAGitty.net.
 50

51 Given the importance of observational studies and attempts to make causal inferences about the impact
 52 of living near AFOs on community health, a thorough evaluation of what is estimated after the control of
 53 unadjusted biasing pathways is vital.
 54

55
 56 **1.2. OBJECTIVES**

57 The first objective will be to document which causal estimate (direct or total) of AFO exposure authors
 58 indicated they were measuring in studies of lower respiratory conditions in communities living in
 59 proximity to AFOs and their rationale for the adjustment set of variables chosen in the study. The
 60 second objective will be to conduct an analysis based on the EPA DAG, to determine which actual effect
 61 sizes were estimated in these studies (direct or total causal effect) and to identify any biasing pathways
 62 remaining after adjustment and what those sources of bias might be.

63 2. Methods

64 2.1. ELIGIBILITY CRITERIA

65 The studies to be used for this report will be obtained from a living systematic review (SR) of
66 epidemiological studies evaluating adverse health outcomes of residents living in areas surrounding
67 AFOs. Studies eligible for the living systematic review were observational studies collecting primary data
68 where the unit of concern for the outcome was the individual. Studies where the unit of measurement
69 of the outcome was a population aggregate (i.e., ecological studies), were not eligible. Participants
70 eligible for inclusion in the review were humans living in communities near AFOs that might be
71 described as industrial, large, concentrated, or other synonyms. Production systems that appeared to be
72 grass-based, nomadic, or confined smallholder operations were also relevant to the review. Exposure to
73 AFOs has been measured in many ways, such as odor intensity, levels of contaminants in the air, soil, or
74 water, proximity measured by distance, or exposure measured by the AFO's animal density units. This
75 list of exposures was indicative rather than exhaustive, and therefore other measures not mentioned
76 previously were also eligible to cover new measures not yet identified. Outcomes of interest were
77 health events or states measured on humans. The outcomes did not need to be a disease; for example,
78 colonization or culture of bacteria from a human is an eligible outcome. Health outcomes captured at a
79 single time, such as self-reported health states or events using survey instruments, were not eligible
80 unless the primary research authors provide evidence of appropriate psychometric properties (validity,
81 reliability, responsiveness) and clinical interpretability (validated), i.e. the authors cited known published
82 disease scales or conditions. Given the wide variety of health outcomes reported in studies, outcomes
83 were categorized by organ system i.e., antimicrobial resistance, gastrointestinal conditions, cancer,
84 infectious conditions, lower respiratory conditions, ocular conditions, psychiatric conditions, skin
85 condition and upper respiratory conditions.

86 The subset of studies from this living systematic review that we will use to meet Objective 1 are the
87 subset of papers that provide comparative estimates of the incidence of health outcomes and
88 residential proximity to AFOs. These papers include cohort studies, incidence case-control studies and
89 prevalence studies that, based on our assessment, meet the population structural assumptions for
90 estimation of causal parameters. Studies that provide comparative estimates of prevalence (either
91 prevalence odds ratios or prevalence ratios) are not relevant in this study, because confounding is a
92 causal concept. For Objective 2, only studies that report lower respiratory disease outcomes or studies
93 that report their own DAG or causal pathway will be considered.

94 2.2. INFORMATION SOURCES

95 The search for this latest group of studies involved a quarterly systematic search methodology, ensuring
96 the identification and assimilation of the most relevant and current studies in the field. The a priori
97 protocol for this review is electronically published with Systematic Reviews for Animals and Food
98 (SYREAF). Electronic searches of MEDLINE®(via Web of Science) (2014 – 2023), CABI Global Health (via
99 Web of Science) (2014 –2023), Centre for Agricultural Biosciences (CAB) Abstracts (via Web of Science)
100 (2014 – 2023), and Science Citation Index (via Web of Science) (2014 – 2023) were conducted.

101 2.3. SEARCH

102 The search strategy used to identify articles on animal feeding operations and community health is
 103 presented in Table 1. The search strategy has two concepts: animal feeding operations and community
 104 health.

105 Table 1. Search strategy in MEDLINE®.

106

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

107

108 *2.4. SELECTION OF SOURCES OF EVIDENCE*

109 Study selection has two levels: a first level based on assessing information in titles and abstracts, and a
 110 second level based on assessing information from the full text of studies. Screening will be conducted
 111 using DistillerSR® (Evidence Partners, Ottawa, ON, Canada). Two independent reviewers will conduct the
 112 selection process (ST, BAFM) and disagreements will be resolved by consulting a third expert reviewer
 113 (AMOC). In the first round of study selection screening, the abstracts and titles will be screened for
 114 eligibility using the following question:

115

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

118

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

122 Is the full text available in English?

123 Is the article a prevalence study i.e. uses a prevalent outcome?

124 Does the study report a comparative association between a relevant animal feeding operation and
 125 measures of health in surrounding-community members?

126 Does the study assess the relationship between outcome and exposure at the individual human level?

127 Does the study report animal feeding operations that would be reasonably considered either large,
 128 concentrated or intensive by modern standards (not nomadic, smallholder or pastoral)?

129 Does the study include more than one unit of measurement of exposure?

130 Does the study include at least one human health outcome measured using either an eligible survey
 131 instrument, test, assay or diseases measure obtained from medical records?

132 For Objective 1 (The targeted causal effect of interest and variable selection approach reported by the
 133 authors) we will further screen the study only to include those that subset of papers that provide
 134 comparative estimates of the incidence of health outcomes and residential proximity to AFOs.

135 For Objective 2, further screening will be for studies that report lower respiratory disease outcomes or
 136 studies that report their own DAG or causal pathway will be considered.

137 *DATA EXTRACTION*

138 For Objective 1, for for each included study, we will assess if the authors reported the target causal
 139 effect, i.e., if the aim of the study was to estimate the direct or total effect of exposure (residential

140 proximity to AFOs) on the health outcomes of interest. Next, we will evaluate the rationale for the
141 adjustment set the authors used in their multivariable models. We will determine if the authors used a
142 multivariable model to obtain an estimate, We will then evaluate the reported rationale for selecting
143 variables for consideration as confounders and, if applicable, the rationale for retention as confounders
144 in the model. For studies that include outcomes from different body systems, we will assess if the
145 authors identified different potential biasing pathways for the different health outcomes.
146 For Objective 2 for each relevant exposure-outcome pair, we will compare the confounding variables
147 included in the model to either a modification of a previously published DAG (Figure 1) proposed to
148 describe the causal association between a lower respiratory condition-chronic bronchitis and living near
149 AFOs or, if available, a DAG provided by the authors. If needed, we will also modify the Brewer et al.
150 2017 DAG to serve as a basis for analyzing and comparing the causal relationships reported by the
151 authors of the primary studies.
152 Subsequently, for each outcome-exposure pair in the relevant papers, we will map the adjustment
153 variables onto either the authors' own proposed DAG or original and any modified Brewer et al. 2017
154 DAG. Using the DAGITTY program, the subsequent DAGs will be evaluated for biasing pathways and if
155 either the total or direct effect causal effect could be estimated. For control variables, we will attempt to
156 match terms that would be consistent, i.e., if the authors adjusted for household income, we will map
157 that to SES and make it an observable variable because the authors implied it was such by adjusting for
158 it. We will then determine if the total effect or the direct effect was estimated and if any biasing
159 pathways remain.

160 *2.5. DATA CHARTING*

161 A data collection form was developed within DistillerSR® to gather relevant data. The form underwent a
162 pretest by two reviewers (BAFM and ST) across the 15 references included until 2014. Subsequently,
163 two reviewers (BAFM and ST) will independently extract the data from all relevant articles utilizing this
164 form. Any discrepancies will be resolved through discussion, and if consensus cannot be reached, a third
165 reviewer will be consulted (AMOC). Information will solely be gathered from the articles themselves; no
166 attempts will be made to contact study investigators for additional or confirmed data. Any missing data
167 will be recorded as 'Not reported', and no assumptions will be made about the unreported information.
168

169 *2.6. DATA ITEMS*

170 We are not extracting data items so this is not relevant.

171 *2.7. CRITICAL APPRAISAL OF INDIVIDUAL SOURCES OF EVIDENCE*

172 As this is a scoping review, we will not be critically appraising the individual sources of evidence.

173 *2.8. SYNTHESIS OF RESULTS*

174 We will use descriptive statistics to summarize our results and DAGs created based on the reported
175 variables controlled for in each of the included studies.
176

177

178

179 References

180

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

183 2. Tricco AC, Lillie E, Zarin W, et al. PRISMA Extension for Scoping Reviews (PRISMA-ScR): Checklist and
184 Explanation. *Ann Intern Med* 2018;169:467-473.

185 3. Brewer LE, Wright JM, Rice G, et al. Causal inference in cumulative risk assessment: The roles of
186 directed acyclic graphs. *Environ Int* 2017;102:30-41.

187