# Developing methods for model selection in causal health analyses: Literature Review Summary

University of Bristol

June 2024

## Plan

#### Queries

Following fiddling / adjustment, we settle on querying the databases with the following :

Web-of-Science:

```
\begin{tabular}{ll} (TI=&(confounder)\ OR\ TI=&(confounding))\ AND\\ (ALL=&(statistical\ model)\ OR\ ALL=&(model\ selection)\ OR\\ ALL=&(variable\ selection))\ AND\ (ALL=&(epidemiology)\ OR\\ ALL=&(causal)) \end{tabular}
```

PubMed:

```
("confounding" [Title/Abstract] OR
"confounder" [Title/Abstract]) AND ("model
selection" [Title/Abstract] OR "variable
selection" [Title/Abstract])
```



# Table of Papers

All papers returned by literature review query:

```
*** All Papers ***

| Familial Confounding of the Association Between Naternal Smoking During Pregnancy and Offspring Substance Use and Problems ... 0
| Assessing seasonal confounding and model selection but an object of the problems of t
```

We have 501 papers from these sources to search through.



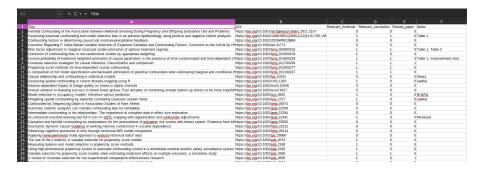
#### **Review Process**

Starting with a set of paper names and DOI's, we assign to each a binary indicator for three categories:

- "Relevant\_methods": contains any novel method or discussion of method behaviour
- "Relevant\_simulation": good examples of simulation studies
- "Review\_papers": papers which are themselves thorough reviews of the literature



#### **Review Process**



## Results

#### Relevant papers:

```
*** Relevant Papers ***

1 Assessing seasonal confounding and model selection bias in air pollution epidemiology using positive and negative control analyses ... 2
4 Each factor adjustment in marginal structural model estimation of optimal treatment regimes ... 2
5 Inverse-probability-of-treatment weighted estimation of causal parameters in the presence of error-contained and time dependent confounders ... 2
7 Covariate selection strategies for causal inference: classification and comparison ... 3
609 Direction dependence analysis A framework to test the direction of effects in linear models with an implementation in 1995 ... 1
609 Confounder detection in linear models and works. Perforance of terror-lawed facts of hodgemoders ... 1
609 Confounder detection in linear models with an implementation in 1995 ... 1
609 Confounder detection in linear models and works. Perforance of terror lawed facts of hodgemodes ... 1
609 SENSITIVITY MALVAISTS FOR MUMERALINE CONFOUNDAME. MOSTED TREAT MODELS ... 1
609 SENSITIVITY MALVAISTS FOR MUMERALINE CONFOUNDAME. MOSTED TREAT MODELS ... 1
609 SEMIPARAMETRIC CAUSAL MEDIATION ANALYSIS WITH UMMEASURED MEDIATOR-OUTCOME COMPOUNDING ... 1
609 SEMIPARAMETRIC CAUSAL MEDIATION ANALYSIS WITH UMMEASURED MEDIATOR-OUTCOME COMPOUNDING ... 1
609 SEMIPARAMETRIC CAUSAL MEDIATION ANALYSIS WITH UMMEASURED MEDIATOR-OUTCOME COMPOUNDING ... 1
609 SEMIPARAMETRIC CAUSAL MEDIATION ANALYSIS WITH UMMEASURED MEDIATOR-OUTCOME COMPOUNDING ... 1
609 SEMIPARAMETRIC CAUSAL MEDIATION ANALYSIS WITH UMMEASURED MEDIATOR-OUTCOME COMPOUNDING ... 1
609 SEMIPARAMETRIC CAUSAL MEDIATION ANALYSIS WITH UMMEASURED MEDIATOR-OUTCOME COMPOUNDING ... 1
609 SEMIPARAMETRIC CAUSAL MEDIATOR ANALYSIS WITH UMMEASURED MEDIATOR-OUTCOME COMPOUNDING ... 1
609 SEMIPARAMETRIC CAUSAL MEDIATION ANALYSIS WITH UMMEASURED MEDIATOR-OUTCOME COMPOUNDING ... 1
609 SEMIPARAMETRIC CAUSAL MEDIATOR ANALYSIS WITH UMMEASURED MEDIATOR-OUTCOME COMPOUNDING ... 1
609 SEMIPARAMETRIC CAUSAL MEDIATION ANALYSIS WITH UMMEASURED MEDIATOR-OUTCOME COMPOUNDING ... 1
```

After removing applications of methods and other irrelevant articles, we have 261 papers left.



#### Relevant methods papers:

Of the 261, 252 contain a relevant discussion of methods.



#### Relevant simulation example papers:

```
Practical recommendations on double score matching for estimating causal effects
                                            Comparison of high-dimensional confounder summary scores in comparative studies of newly marketed medications
                                                                                                      SIMULATION STUDY OF CONFOUNDER-SELECTION STRATEGIES
                                                                               The change in estimate method for selecting confounders: A simulation study
414 Comparing q-computation, propensity score-based weighting, and targeted maximum likelihood estimation for analyzing externally controlled trials w...
```

Of the 261, 35 are good simulation examples worth emulating.





#### Key review papers:

```
Covariate selection strategies for causal inference: Classification and comparative of covariate selection strategies for causal inference: Classification and comparative and comparative affects of covariate selection for non-experimental comparative affects sense and comparative affects of covariate selection for non-experimental comparative affects sense sense and comparative affects of covariate selection and adjustment methods for estimating causal effects using large healthcare databases as a sense of controlled adjustment and configuration of controlled studies are controlled in the controlled studies as a sense of controlled studies are controlled studies as a sense of controlled studies are controlled studies as a sense of controlled studies are controlled studies are controlled studies as a sense of controlled studies are controlled and sense of controlled studies are controlled and sense of controlled studies are controlled and sense of controlled studies are controlled studies and sense are controlled studies are controlled studies. A comparison of Methods to Estimate the loader conditions on the controlled studies are controlled studies. A comparison of Methods to Estimate the loader conditions of management and studies are controlled studies. A comparison of Methods to Estimate the loader conditions of water controlled studies are controlled studies. A comparison of Methods to Estimate the loader conditions of the controlled studies are controlled studies. A comparison of Methods to Estimate the loader conditions of the controlled studies are controlled studies. A comparison of Methods to Est
```

Of the 261, 21 are themselves key reviews of the literature.



#### "Best" papers:

```
*** Relevant Best Papers ***

Covariate selection strategies for causal inference: Classification and comparison ... 3

A comparison of confounder selection and adjustment methods for estimating causal effects using large healthcare databases ... 3

414 Comparing g-computation, propensity score-based weighting, and targeted maximum tikelihood estimation for analyzing externally controlled trials w... 3

[3 rows x 7 columns]
```

Of the 261, 3 papers were relevant in all three categories simultaneously.

#### Methods List

From this, we move towards a list of all 115 methods encountered in the literature with claims around confounder handling:

