



MRC Integrative
Epidemiology
Unit



University of
BRISTOL

qbaconfound: A flexible Monte Carlo probabilistic bias analysis for unmeasured confounding

Emily Kawabata, Chin Yang Shapland, Tom Palmer,
David Carslake, Kate Tilling, Rachael Hughes

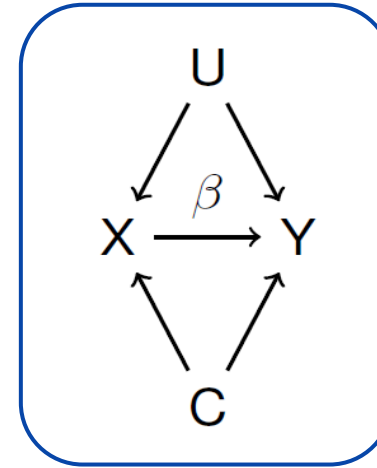
Population Health Sciences, Bristol Medical School, University of Bristol
MRC Integrative Epidemiology Unit, University of Bristol

Supported by a Sir Henry Dale Fellowship
jointly funded by the Wellcome Trust and
the Royal Society, and by the MRC



Accounting for unmeasured confounding

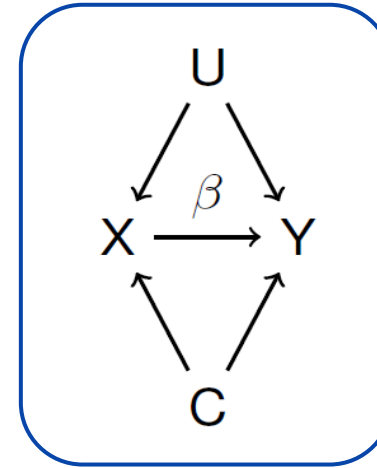
- Aim is to estimate causal exposure effect
 - exposure X
 - outcome Y
 - causal exposure effect β
 - measured confounders C (of $X \rightarrow Y$)
 - unmeasured confounders U (of $X \rightarrow Y$)



- Conventional analysis adjusts for C assumes no unmeasured confounding
- Failure to account for unmeasured or poorly measured confounders may lead to invalid inference
- Several approaches to assess causality such as:
 - self-controlled study design, instrumental variable analysis, negative controls, perturbation variable analysis, quantitative bias analysis

Accounting for unmeasured confounding

- Aim is to estimate causal exposure effect
 - exposure X
 - outcome Y
 - causal exposure effect β
 - measured confounders C (of $X \rightarrow Y$)
 - unmeasured confounders U (of $X \rightarrow Y$)



- Conventional analysis adjusts for C assumes no unmeasured confounding
- Failure to account for unmeasured or poorly measured confounders may lead to invalid inference
- Several approaches to assess causality such as:
 - self-controlled study design, instrumental variable analysis, negative controls, perturbation variable analysis, **quantitative bias analysis**

Quantitative bias analysis (QBA)

- QBA can be used to quantify
 - impact of unmeasured confounding on exposure effect,
 - level of unmeasured confounding required to change conclusions
- QBA requires a bias model with bias parameters θ
 - e.g., specify strength of the association between U and X given C , and between U and Y given X and C
 - generate bias-adjusted estimate of exposure effect, β
- Information about θ may be obtained from
 - external sources (e.g., published literature, data from another study)
 - measured confounders C (known as benchmarking or calibration)
- Two broad classes of QBA methods: deterministic and probabilistic

Probabilistic QBA

- Prior distribution $p(\theta)$ to incorporate analyst's assumptions about
 - plausible values for θ
 - value-combinations that are most likely to occur
 - uncertainty about the values of θ
- Using $p(\theta)$ and data on (Y, X, C) , generate distribution of bias-adjusted estimates of exposure effect, β
 - summarise distribution: point estimate and interval estimates
- Bayesian QBA: combines $p(\theta)$ with likelihood function for the data
- Monte Carlo QBA: Samples directly from $p(\theta)$ then fixes θ to sampled values to give bias-adjusted estimate of β . Repeats many times.

RESEARCH

Open Access

Quantitative bias analysis in practice: review of software for regression with unmeasured confounding



Emily Kawabata^{1,2*}, Kate Tilling^{1,2}, Rolf H. H. Groenwold^{3,4} and Rachael A. Hughes^{1,2}

- Programs tended to be specific to one type of analysis
- Wide diversity in the types of QBA method implemented

Existing software implementing a probabilistic QBA

- Software implementing a probabilistic QBA include
 - Stata command *episens*, R packages *CLMTx*, *episensr*, *SAMTx*, and *unmconf* and R Shiny application *apisensr*
- R package *unmconf* (Hebdon et al 2024)
 - implements a Bayesian QBA
 - applicable when the analysis is a generalised linear model (GLM)
 - allows for multiple continuous or binary unmeasured confounders
 - does not assume independence between any of the confounders
 - external information about θ : validation data
 - for an analysis with a binary or continuous exposure:
 - 13 to 14 bias parameters for 10 measured confounders and 1 unmeasured confounder
 - 27 to 29 bias parameters for 10 measured confounders and 2 unmeasured confounders

Our proposed probabilistic QBA

- Flexible probabilistic QBA tool when only summary data are available on θ
- Applicable when analysis of interest is a
 - generalised linear model
 - survival proportional hazards regression
- Allows for multiple continuous or binary unmeasured confounders
- Does not assume independence between any of the confounders
- Requires a smaller number of bias parameters θ
- For an analysis with a binary or continuous exposure:
 - 3 bias parameters for 10 measured confounders and 1 unmeasured confounders
 - 6 bias parameters for 10 measured confounders and 2 unmeasured confounders
- Implemented as a Monte Carlo QBA

Illustrative example

- National Health and Nutritional Examination Survey
- Based on a study by Butler et al (2018)
- Y is waist circumference (cm)
- $X = (X_{nev}, X_{form}, X_{3to4}, X_{5+})$ is average alcohol consumption in one drinking day
 - never (X_{nev}), former (X_{form}), 1-2 drinks, 3-4 drinks (X_{3to4}), ≥ 5 drinks (X_{5+})
- C are age, ethnicity, education, income, survey year, smoking status, day of dietary recall, physical activity level, dietary misreporting status
- Continuous U_1 and binary U_2
 - U_1 is daily non-alcoholic energy intake
 - U_2 is marital status (0 currently married; 1 otherwise)
 - measured by NHANES; Butler et al adjusted for U_1 and U_2



- Analysis of interest

$$Y|X, C, U \sim N(\beta_0 + \beta_X X^T + \beta_C C^T + \beta_{U_1} U_1 + \beta_{U_2} U_2, \sigma^2)$$

- $\beta_X = (\beta_{X_{nev}}, \beta_{X_{form}}, \beta_{X_{3to4}}, \beta_{X_{5+}})$

Rationale underlying proposed bias model

- Analysis of interest

$$Y|X, C, U \sim N(\beta_0 + \beta_X X^T + \beta_C C^T + \beta_{U_1} U_1 + \beta_{U_2} U_2, \sigma^2)$$

- $\beta_X = (\beta_{X_{nev}}, \beta_{X_{form}}, \beta_{X_{3to4}}, \beta_{X_{5+}})$

- Rationale

- simulate U_1 and U_2
- fit analysis constraining coefficients of U_1 and U_2 to β_{U_1} and β_{U_2} , respectively

$$U_1|X, C \sim N(\alpha_{1,X} X^T + \alpha_{1,C} C^T, \eta_1^2)$$

$$U_2|X, C \sim \text{Bern}(\text{invlogit}\{\alpha_{2,0} + \alpha_{2,X} X^T + \alpha_{2,C} C^T\})$$

$$Y|X, C, U_1, U_2 \sim N(\beta_0 + \beta_X X^T + \beta_C C^T + \beta_{U_1} U_1 + \beta_{U_2} U_2, \sigma^2)$$

Rationale underlying proposed bias model

- Analysis of interest

$$Y|X, C, U \sim N(\beta_0 + \beta_X X^T + \beta_C C^T + \beta_{U_1} U_1 + \beta_{U_2} U_2, \sigma^2)$$

- $\beta_X = (\beta_{X_{nev}}, \beta_{X_{form}}, \beta_{X_{3to4}}, \beta_{X_{5+}})$

- Rationale

- simulate U_1 and U_2
- fit analysis constraining coefficients of U_1 and U_2 to β_{U_1} and β_{U_2} , respectively

$$U_1|X, C \sim N(\alpha_{1,X} X^T + \alpha_{1,C} C^T, \eta_1^2)$$

$$U_2|X, C \sim \text{Bern}(\text{invlogit}\{\alpha_{2,0} + \alpha_{2,X} X^T + \alpha_{2,C} C^T\})$$

$$Y|X, C, U_1, U_2 \sim N(\beta_0 + \beta_X X^T + \beta_C C^T + \beta_{U_1} U_1 + \beta_{U_2} U_2, \sigma^2)$$

Proposed bias model

- Simulate $\tilde{U}_1$: the part of U_1 uncorrelated with C
- Simulate $\tilde{U}_2$: the part of U_2 uncorrelated with C
- Fit analysis constraining coefficients of $\tilde{U}_1$ and $\tilde{U}_2$ to β_{U_1} and β_{U_2} , respectively

$$\tilde{U}_1 | X, C \sim N(\alpha_{1,X} X^T, \eta_1^2)$$

$$\tilde{U}_2 | X, C \sim \text{Bern}(\text{invlogit}\{\tilde{\alpha}_{2,0} + \alpha_{2,X} X^T\})$$

$$Y | X, C, \tilde{U}_1, \tilde{U}_2 \sim N(\tilde{\beta}_0 + \beta_X X^T + \tilde{\beta}_C C^T + \beta_{U_1} \tilde{U}_1 + \beta_{U_2} \tilde{U}_2, \sigma^2)$$

Proposed bias model

- Simulate $\tilde{U}_1$: the part of U_1 uncorrelated with C
- Simulate $\tilde{U}_2$: the part of U_2 uncorrelated with C
- Fit analysis constraining coefficients of $\tilde{U}_1$ and $\tilde{U}_2$ to β_{U_1} and β_{U_2} , respectively

$$\tilde{U}_1 | X, C \sim N(\alpha_{1,X} X^T, \eta_1^2)$$

$$\tilde{U}_2 | X, C \sim \text{Bern}(\text{invlogit}\{\tilde{\alpha}_{2,0} + \alpha_{2,X} X^T\})$$

$$Y | X, C, \tilde{U}_1, \tilde{U}_2 \sim N(\tilde{\beta}_0 + \beta_X X^T + \tilde{\beta}_C C^T + \beta_{U_1} \tilde{U}_1 + \beta_{U_2} \tilde{U}_2, \sigma^2)$$

- $\pi_2 = \text{Pr}(U_2 = 1)$
 - derive $\tilde{\alpha}_{2,0}$ from π_2 and $\alpha_{2,X}$ using a numerical approximation approach
- Bias parameters $\theta = (\beta_{U_1}, \beta_{U_2}, \alpha_{1,X}, \eta_1, \alpha_{2,X}, \pi_2)$

Proposed bias model

- Simulate $\tilde{U}_1$: the part of U_1 uncorrelated with C
- Simulate $\tilde{U}_2$: the part of U_2 uncorrelated with C
- Fit analysis constraining coefficients of $\tilde{U}_1$ and $\tilde{U}_2$ to β_{U_1} and β_{U_2} , respectively

$$\tilde{U}_1 | X, C \sim N(\alpha_{1,X} X^T, \eta_1^2)$$

$$\tilde{U}_2 | X, C \sim \text{Bern}(\text{invlogit}\{\tilde{\alpha}_{2,0} + \alpha_{2,X} X^T\})$$

$$Y | X, C, \tilde{U}_1, \tilde{U}_2 \sim N(\tilde{\beta}_0 + \beta_X X^T + \tilde{\beta}_C C^T + \beta_{U_1} \tilde{U}_1 + \beta_{U_2} \tilde{U}_2, \sigma^2)$$

- $\pi_2 = \text{Pr}(U_2 = 1)$
 - derive $\tilde{\alpha}_{2,0}$ from π_2 and $\alpha_{2,X}$ using a numerical approximation approach
- Bias parameters $\theta = (\beta_{U_1}, \beta_{U_2}, \alpha_{1,X}, \eta_1, \alpha_{2,X}, \pi_2)$

Independent prior distributions for θ

$$p(\theta) = p(\beta_{U_1})p(\beta_2)p(\alpha_{1,X_{nev}}) \dots p(\alpha_{1,X_{5+}})p(\eta_1)p(\alpha_{2,X_{nev}}) \dots p(\alpha_{2,X_{5+}})p(\pi_2)$$

- Normal distribution for each coefficient
 - e.g., $p(\alpha_{1,X_{5+}}) \sim N(\mu_{\alpha_{1,X_{5+}}}, \psi_{\alpha_{1,X_{5+}}})$
- Uniform distribution for
 - standard deviation term: $p(\eta_1) \sim Un(Lower_{\eta_1}, Upper_{\eta_1})$
 - prevalence: $p(\pi_2) \sim Un(Lower_{\pi_2}, Upper_{\pi_2})$
- Set hyperparameters to observed data estimates; e.g., fit $U_1 | X, C$
 - coefficient estimate for $\mu_{\alpha_{1,X_{5+}}} = -94.71$
 - accompanying standard error for $\psi_{\alpha_{1,X_{5+}}} = 19.64$

Details on the proposed method

- Accepted for publication in Statistical Methods in Medical Research
- Preprint available below

medRxiv

THE PREPRINT SERVER FOR HEALTH SCIENCES

qbaconfound: A flexible Monte Carlo probabilistic bias analysis for unmeasured confounding

 Emily Kawabata,  Chin Yang Shapland,  Tom M Palmer, David Carslake, Kate Tilling, Rachael Hughes

doi: <https://doi.org/10.1101/2025.08.12.25333217>

Syntax for one or two unmeasured confounders

```
qbaconfound depvar indepvars [if] [in] , treatment(treatvar) ///  
    firstuvar( [newvar] , coefuonx(numlist) sdcoefuonx(numlist) ///  
                loweruni(numlist) upperuni(numlist) ///  
                coefyonu(numlist) sdcoefyonu(numlist) [isbinary]) ///  
[ seconduvar(...) other_options ]
```

- Required options

- treatment(*treatvars*) identifies the treatment variable of interest
- firstuvar(...) provides information about one unmeasured confounder

- By default

- analysis model is a linear regression and includes a constant term
- one continuous unmeasured confounder
- 100 Monte Carlo replications

Syntax for one or two unmeasured confounders

```
qbaconfound depvar indepvars [if] [in] , treatment(treatvar) ///  
    firstuvar( [newvar] , coefuonx(numlist) sdcoefuonx(numlist) ///  
                loweruni(numlist) upperuni(numlist) ///  
                coefyonu(numlist) sdcoefyonu(numlist) [isbinary]) ///  
[ seconduvar(...) other_options ]
```

- Required options

- treatment(*treatvars*) identifies the treatment variable of interest
- firstuvar(...) provides information about one unmeasured confounder

- By default

- analysis model is a linear regression and includes a constant term
- one continuous unmeasured confounder
- 100 Monte Carlo replications

Syntax for one or two unmeasured confounders

```
qbaconfound depvar indepvars [if] [in] , treatment(treatvar) ///  
    firstuvar( [newvar] , coefuonx(numlist) sdcoefuonx(numlist) ///  
        loweruni(numlist) upperuni(numlist) ///  
        coefyonu(numlist) sdcoefyonu(numlist) [isbinary]) ///  
    [ seconduvar(...) other_options ]
```

- Required options

- treatment(*treatvars*) identifies the treatment variable of interest
- firstuvar(...) provides information about one unmeasured confounder

- By default

- analysis model is a linear regression and includes a constant term
- one continuous unmeasured confounder
- 100 Monte Carlo replications

Syntax for one or two unmeasured confounders

```
qbaconfound depvar indepvars [if] [in] , treatment(treatvar) ///  
    firstuvar( [newvar] , coefuonx(numlist) sdcoefuonx(numlist) ///  
                loweruni(numlist) upperuni(numlist) ///  
                coefyonu(numlist) sdcoefyonu(numlist) [isbinary]) ///  
[ seconduvar(...) other_options ]
```

- Required options

- treatment(*treatvars*) identifies the treatment variable of interest
- firstuvar(...) provides information about one unmeasured confounder

- By default

- analysis model is a linear regression and includes a constant term
- one continuous unmeasured confounder
- 100 Monte Carlo replications

Syntax for one or more unmeasured confounders

```
qbaconfound depvar indepvars [if] [in] , treatment(treatvars) ///  
    multiuvar([stub] , coefmatrix(name) sdcoefmatrix(name) ///  
        unimatrix(name) [isbinarymatrix(name)] ///  
    [ other_options ]
```

- Required options

- treatment(*treatvars*) identifies the treatment variable of interest
- multiuvar(...) provides information about unmeasured confounder(s)

- By default

- analysis model is a linear regression and includes a constant term
- one continuous unmeasured confounder
- 100 Monte Carlo replications

Options

- Analysis of interest
 - `command(name)`
 - *name*: regress, logit, logistic, mlogit, ologit, stcox
 - `noconstant`
 - `eform`
- Monte Carlo QBA
 - `dryrun`
 - `reps(integer)`
 - `seed(integer)`
 - `savemcreps(filename, replace)`
- Graphical options
 - `grprior(name)` histogram of the bias parameters values drawn from specified priors
 - `grsyserr(name)` histogram of the bias-adjusted estimates (systemic error)
 - `grtoterr(name)` histogram of the bias-adjusted estimates (systemic plus random error)

Example - dryrun

```
qbaconfound waist ib1.alcohol ///
  i.(age race educ physical year recall* smoking income misreport), ///
  treat(ib1.alcohol) dryrun ///
  firstuvar(energy, coefuonx(-65.37 -94.71 69.72 -43.47) ///
    sdcoefuonx(18.07 19.64 18.63 27.21) ///
    coefyonu(0.006931) sdcoefyonu(0.0002805) ///
    loweruni(372.6) upperuni(716.5)) ///
  seconduvar(marital, coefuonx(0.3831 0.3785 -0.07963 0.4524) ///
    sdcoefuonx(0.07405 0.07958 0.08107 0.1081) ///
    coefyonu(-1.312) sdcoefyonu(0.3453) ///
    loweruni(0.2900) upperuni(0.3484) isbinary)
```

Example - dryrun

```
qbaconfound waist ib1.alcohol ///  
  i.(age race educ physical year recall* smoking income misreport), ///  
  treat(ib1.alcohol) dryrun ///  
firstuvar(energy, coefuonx(-65.37 -94.71 69.72 -43.47) ///  
          sdcoefuonx(18.07 19.64 18.63 27.21) ///  
          coefyonu(0.006931) sdcoefyonu(0.0002805) ///  
          loweruni(372.6) upperuni(716.5)) ///  
seconduvar(marital, coefuonx(0.3831 0.3785 -0.07963 0.4524) ///  
             sdcoefuonx(0.07405 0.07958 0.08107 0.1081) ///  
             coefyonu(-1.312) sdcoefyonu(0.3453) ///  
             loweruni(0.2900) upperuni(0.3484) isbinary)
```

Example - dryrun

```
qbaconfound waist ib1.alcohol ///
  i.(age race educ physical year recall* smoking income misreport), ///
  treat(ib1.alcohol) dryrun ///
firstuvar(energy, coefuonx(-65.37 -94.71 69.72 -43.47) ///
sdcoefuonx(18.07 19.64 18.63 27.21) ///
coefyonu(0.006931) sdcoefyonu(0.0002805) ///
loweruni(372.6) upperuni(716.5)) ///
seconduvar(marital, coefuonx(0.3831 0.3785 -0.07963 0.4524) ///
  sdcoefuonx(0.07405 0.07958 0.08107 0.1081) ///
  coefyonu(-1.312) sdcoefyonu(0.3453) ///
  loweruni(0.2900) upperuni(0.3484) isbinary)
```

Example - dryrun

```
qbaconfound waist ib1.alcohol ///
  i.(age race educ physical year recall* smoking income misreport), ///
  treat(ib1.alcohol) dryrun ///
  firstuvar(energy, coefuonx(-65.37 -94.71 69.72 -43.47) ///
    sdcoefuonx(18.07 19.64 18.63 27.21) ///
    coefyonu(0.006931) sdcoefyonu(0.0002805) ///
    loweruni(372.6) upperuni(716.5)) ///
seconduvar(marital, coefuonx(0.3831 0.3785 -0.07963 0.4524) ///
  sdcoefuonx(0.07405 0.07958 0.08107 0.1081) ///
  coefyonu(-1.312) sdcoefyonu(0.3453) ///
  loweruni(0.2900) upperuni(0.3484) isbinary)
```

Example - dryrun

```
qbaconfound waist ib1.alcohol ///
  i.(age race educ physical year recall* smoking income misreport), ///
  treat(ib1.alcohol) dryrun ///
  firstuvar(energy, coefuonx(-65.37 -94.71 69.72 -43.47) ///
    sdcoefuonx(18.07 19.64 18.63 27.21) ///
    coefyonu(0.006931) sdcoefyonu(0.0002805) ///
    loweruni(372.6) upperuni(716.5)) ///
  seconduvar(marital, coefuonx(0.3831 0.3785 -0.07963 0.4524) ///
    sdcoefuonx(0.07405 0.07958 0.08107 0.1081) ///
    coefyonu(-1.312) sdcoefyonu(0.3453) ///
    loweruni(0.2900) upperuni(0.3484) isbinary)
```

Prior distributions specified; no results generated

Model	Bias parameter	Prior distribution
regress energy i.alco~1 C	2.alco~1 3.alco~1 4.alco~1 5.alco~1 Residual SD	Normal(-65.37, 18.07) Normal(-94.70999999999999, 19.64) Normal(69.72, 18.63) Normal(-43.47, 27.21) Uniform(372.6, 716.5)
logit marital i.alco~1 C	2.alco~1 3.alco~1 4.alco~1 5.alco~1	Normal(.3831, .07405) Normal(.3785, .07958) Normal(-.07963, .08107) Normal(.4524, .1081)
marital	Proportion	Uniform(.29, .3484)
regress waistc~y i.alco~1 C U	energy marital	Normal(.006931, .0002805) Normal(-1.312, .3453)

C: measured confounders, U: unmeasured confounders

Example – conducts probabilistic bias analysis

```
qbaconfound waist ib1.alcohol ///
  i.(age race educ physical year recall* smoking income misreport), ///
  treat(ib1.alcohol) reps(10000) seed(987) ///
  firstuvar(energy, coefuonx(-65.37 -94.71 69.72 -43.47) ///
    sdcoefuonx(18.07 19.64 18.63 27.21) ///
    coefyonu(0.006931) sdcoefyonu(0.0002805) ///
    loweruni(372.6) upperuni(716.5)) ///
  seconduvar(marital, coefuonx(0.3831 0.3785 -0.07963 0.4524) ///
    sdcoefuonx(0.07405 0.07958 0.08107 0.1081) ///
    coefyonu(-1.312) sdcoefyonu(0.3453) ///
    loweruni(0.2900) upperuni(0.3484) isbinary) ///
grprior("bpdraws") grsyserr("appsyserr") grtoterr("apptoterr")
```

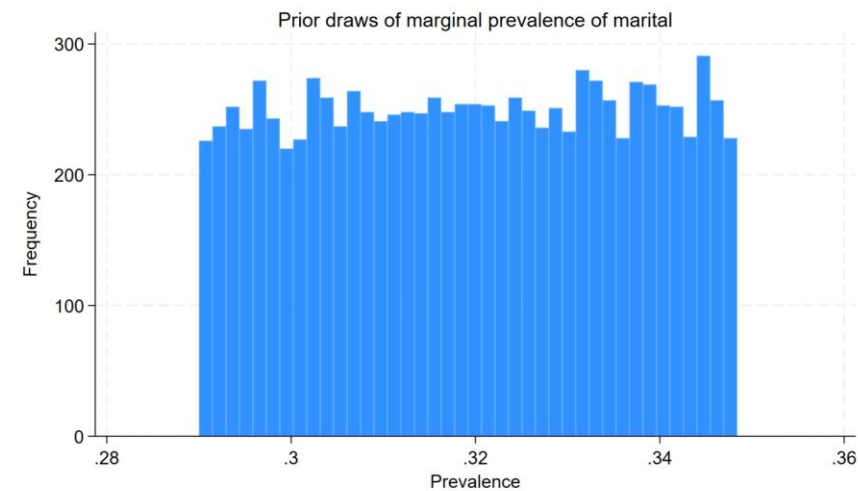
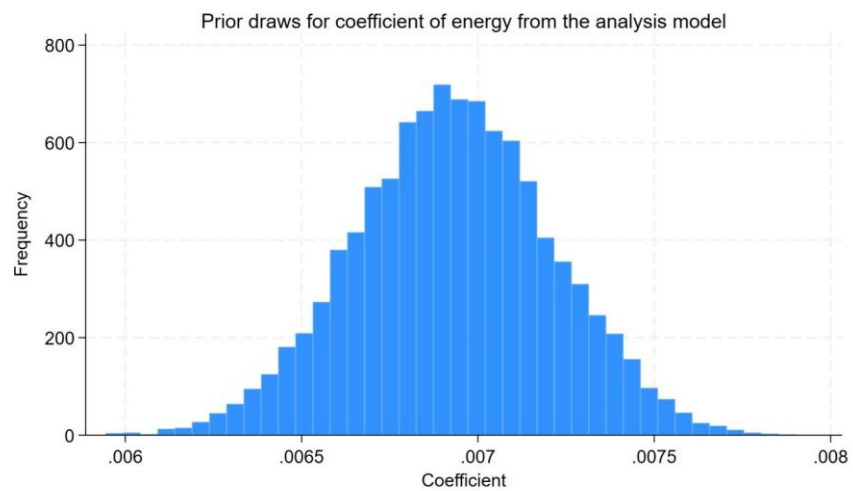
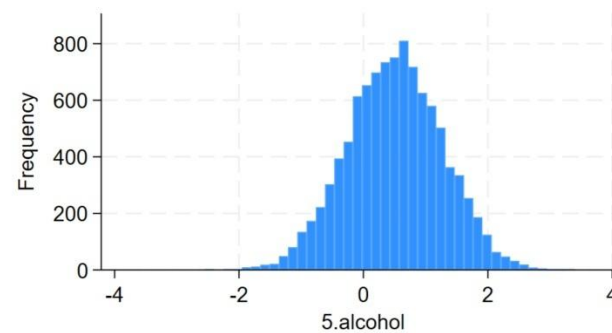
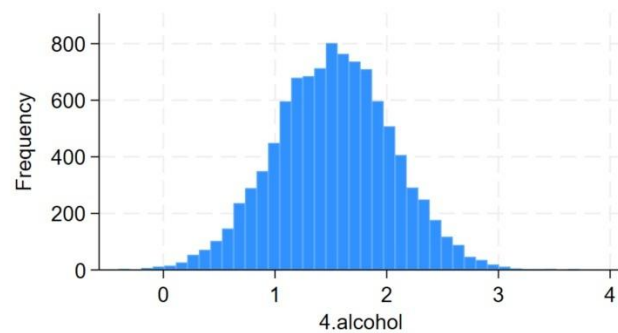
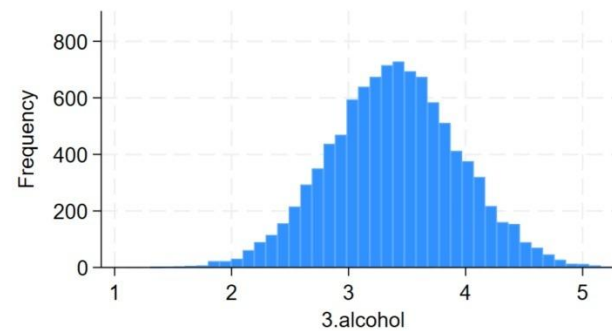
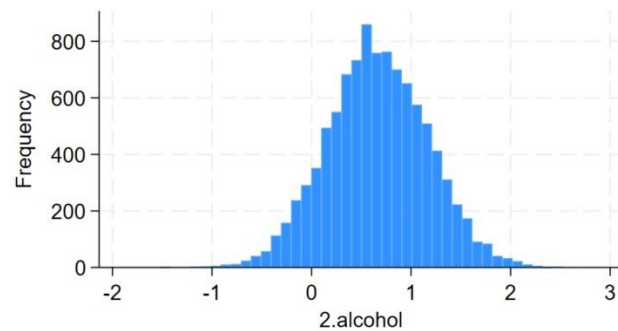
Probabilistic bias analysis results

Number of reps = 10000

Number of obs = 7477

waistcircumferencey	Coefficient	[95% Interval]	
2.alcohol:			
conventional (assumes no bias)	.1027047	-.7899326	.9953419
systematic error only	.6625046	.3044252	1.033908
random and systematic error	.6533227	-.3223721	1.662631
3.alcohol:			
conventional (assumes no bias)	2.621148	1.651198	3.591098
systematic error only	3.385628	2.989425	3.783769
random and systematic error	3.389882	2.295886	4.482009
4.alcohol:			
conventional (assumes no bias)	2.037006	1.116571	2.957441
systematic error only	1.535626	1.167242	1.904974
random and systematic error	1.531037	.4877835	2.557544
5.alcohol:			
conventional (assumes no bias)	.0696549	-1.274649	1.413958
systematic error only	.5013274	-.0430503	1.04666
random and systematic error	.5106779	-.9978893	1.978951

Bias adjusted estimates: random and systematic error



Final remarks

- Plan to submit an article to Stata Journal
- Also plan to develop corresponding R package *qbaconfound*
- Sourcing external information about the bias parameters is challenging
 - bias parameters expressed with respect to the original unmeasured confounders
- Further work
 - implement as a deterministic bias analysis
 - investigate application of *qbaconfound* in other settings
 - e.g., weighted analysis, multilevel analysis, meta-analysis
 - how to deal with missing data

References

- Butler B, Popkin BM, Poti JM, “Associations of alcoholic beverage consumption with dietary intake, waist circumference, and body mass index in US adults: National Health and Nutrition Examination Survey 2003-2012,” *J Acad Nutr Diet*, 2018; 118: 409-20.
- Hebdon R, Stamey J, Kahle D, Zhang X. “unmconf: an R package for Bayesian regression with unmeasured confounders,” *BMC Med Res Methodol*, 2024; 24:195.
- R. Hebdon, J. Stamey, D. Kahle and X. Zhang, “How to Use the unmconf Package,” [Online]. Available: <https://cran.r-project.org/web/packages/unmconf/vignettes/unmconf.html>
- Kawabata E, Tilling K, Groenwold R, Hughes R. “Quantitative bias analysis in practice: review of software for regression with unmeasured confounding”. *BMC Med Res Methodol*, 2023; 23:111.



MRC Integrative
Epidemiology
Unit



University of
BRISTOL

Thank you

Monte Carlo QBA using our proposed bias model

For $k = 1, \dots, K$

1. Randomly draw $\theta^{(k)} \sim p(\theta)$
 - derive $\tilde{\alpha}_{2,0}^{(k)}$ from drawn values $\pi_2^{(k)}$ and $\alpha_{2,X}^{(k)}$
 2. Fit the bias model to (Y, X, C) , fixing bias parameters to drawn values $\theta^{(k)}$
 - a. simulate $\tilde{U}_1^{(k)} \sim N\left(\alpha_{1,X}^{(k)} X^T, \eta_1^{(k)^2}\right)$
 - b. simulate $\tilde{U}_2^{(k)} \sim \text{Bern}\left(\text{invlogit}\left\{\tilde{\alpha}_{2,0}^{(k)} + \alpha_{2,X}^{(k)} X^T\right\}\right)$
 - c. fit constrained linear regression $\left(\tilde{\beta}_0 + \beta_X X^T + \tilde{\beta}_C C^T + \beta_{U_1}^{(k)} \tilde{U}_1^{(k)} + \beta_{U_2}^{(k)} \tilde{U}_2^{(k)}, \sigma^2\right)$
 3. Incorporate random sampling error
 - e.g., $\beta_{X_{5+}}^{(k)} \sim N(\hat{\beta}_{X_{5+}}, \widehat{SE}_{X_{5+}})$
- Summarise empirical distribution of each exposure effect (e.g., $\beta_{X_{5+}}^{(1)}, \beta_{X_{5+}}^{(2)}, \dots, \beta_{X_{5+}}^{(K)}$) as median and [2.5th, 97.5th percentiles]