Population unattributable fractions and other scenario comparisons

With examples from the Avon Longitudinal Study of Parents and Children (ALSPAC) cohort study at Bristol University, UK
http://www.bristol.ac.uk/alspac/

Roger B. Newson
r.newson@imperial.ac.uk
http://www.imperial.ac.uk/nhli/r.newson/

National Heart and Lung Institute
Imperial College London

Asthma Club, 11 November, 2010

Why scenario comparisons?

- Public health scientists make their living mostly by proposing (or fantasizing) **interventions**.
- Such interventions might be helping people to quit smoking, offering informative DNA tests, or genetic engineering of eggs and sperms.
- A skeptical public will ask what good these proposed interventions will do, especially if they are expensive.
- *So* public health scientists need to be able to give an answer.
- These answers are comparisons between **scenarios**, defined as fantasy worlds in which different interventions are made (or not made).
- Any two fantasy scenarios compared will be different in some ways, but the same in others.

Why scenario comparisons?

- Public health scientists make their living mostly by proposing (or fantasizing) **interventions**.
- Such interventions might be helping people to quit smoking, offering informative DNA tests, or genetic engineering of eggs and sperms.
- A skeptical public will ask what good these proposed interventions will do, especially if they are expensive.
- *So* public health scientists need to be able to give an answer.
- These answers are comparisons between **scenarios**, defined as fantasy worlds in which different interventions are made (or not made).
- Any two fantasy scenarios compared will be different in some ways, but the same in others.

Why scenario comparisons?

- Public health scientists make their living mostly by proposing (or fantasizing) **interventions**.
- Such interventions might be helping people to quit smoking, offering informative DNA tests, or genetic engineering of eggs and sperms.
- A skeptical public will ask what good these proposed interventions will do, especially if they are expensive.
- *So* public health scientists need to be able to give an answer.
- These answers are comparisons between **scenarios**, defined as fantasy worlds in which different interventions are made (or not made).
- Any two fantasy scenarios compared will be different in some ways, but the same in others.

Why scenario comparisons?

- Public health scientists make their living mostly by proposing (or fantasizing) **interventions**.
- Such interventions might be helping people to quit smoking, offering informative DNA tests, or genetic engineering of eggs and sperms.
- A skeptical public will ask what good these proposed interventions will do, especially if they are expensive.
- *So* public health scientists need to be able to give an answer.
- These answers are comparisons between **scenarios**, defined as fantasy worlds in which different interventions are made (or not made).
- Any two fantasy scenarios compared will be different in some ways, but the same in others.

Why scenario comparisons?

- Public health scientists make their living mostly by proposing (or fantasizing) **interventions**.
- Such interventions might be helping people to quit smoking, offering informative DNA tests, or genetic engineering of eggs and sperms.
- A skeptical public will ask what good these proposed interventions will do, especially if they are expensive.
- *So* public health scientists need to be able to give an answer.
- These answers are comparisons between **scenarios**, defined as fantasy worlds in which different interventions are made (or not made).
- Any two fantasy scenarios compared will be different in some ways, but the same in others.

Why scenario comparisons?

- Public health scientists make their living mostly by proposing (or fantasizing) **interventions**.
- Such interventions might be helping people to quit smoking, offering informative DNA tests, or genetic engineering of eggs and sperms.
- A skeptical public will ask what good these proposed interventions will do, especially if they are expensive.
- *So* public health scientists need to be able to give an answer.
- These answers are comparisons between **scenarios**, defined as fantasy worlds in which different interventions are made (or not made).
- Any two fantasy scenarios compared will be different in some ways, but the same in others.

Why scenario comparisons?

- Public health scientists make their living mostly by proposing (or fantasizing) **interventions**.
- Such interventions might be helping people to quit smoking, offering informative DNA tests, or genetic engineering of eggs and sperms.
- A skeptical public will ask what good these proposed interventions will do, especially if they are expensive.
- *So* public health scientists need to be able to give an answer.
- These answers are comparisons between **scenarios**, defined as fantasy worlds in which different interventions are made (or not made).
- Any two fantasy scenarios compared will be different in some ways, but the same in others.

What are scenario comparison statistics?

- In statistics, **scenarios** can be defined as alternative versions of the same dataset.
- *For instance*, we might have a dataset with 1 observation per patient, and data on age, smoking and lung disease.
- And we might compute an alternative version of the same dataset, with the same age distribution, but all patients lifelong non–smokers.
- Lung disease outcome distributions in these two versions can be **compared** using a **statistic**, such as a mean difference, a mean ratio, a ratio between means, a median difference, or a difference between medians.
- And all of these comparisons can be presented with confidence limits and *P*–values.
- We can therefore estimate the consequences of a proposed intervention, *assuming* that an association between smoking and lung disease is **causal**, at least within each age group.

What are scenario comparison statistics?

- In statistics, **scenarios** can be defined as alternative versions of the same dataset.
- *For instance*, we might have a dataset with 1 observation per patient, and data on age, smoking and lung disease.
- And we might compute an alternative version of the same dataset, with the same age distribution, but all patients lifelong non–smokers.
- Lung disease outcome distributions in these two versions can be **compared** using a **statistic**, such as a mean difference, a mean ratio, a ratio between means, a median difference, or a difference between medians.
- And all of these comparisons can be presented with confidence limits and *P*–values.
- We can therefore estimate the consequences of a proposed intervention, *assuming* that an association between smoking and lung disease is **causal**, at least within each age group.

What are scenario comparison statistics?

- In statistics, **scenarios** can be defined as alternative versions of the same dataset.
- *For instance*, we might have a dataset with 1 observation per patient, and data on age, smoking and lung disease.
- And we might compute an alternative version of the same dataset, with the same age distribution, but all patients lifelong non–smokers.
- Lung disease outcome distributions in these two versions can be **compared** using a **statistic**, such as a mean difference, a mean ratio, a ratio between means, a median difference, or a difference between medians.
- And all of these comparisons can be presented with confidence limits and *P*–values.
- We can therefore estimate the consequences of a proposed intervention, *assuming* that an association between smoking and lung disease is **causal**, at least within each age group.

What are scenario comparison statistics?

- In statistics, **scenarios** can be defined as alternative versions of the same dataset.
- *For instance*, we might have a dataset with 1 observation per patient, and data on age, smoking and lung disease.
- And we might compute an alternative version of the same dataset, with the same age distribution, but all patients lifelong non–smokers.
- Lung disease outcome distributions in these two versions can be **compared** using a **statistic**, such as a mean difference, a mean ratio, a ratio between means, a median difference, or a difference between medians.
- And all of these comparisons can be presented with confidence limits and *P*–values.
- We can therefore estimate the consequences of a proposed intervention, *assuming* that an association between smoking and lung disease is **causal**, at least within each age group.

What are scenario comparison statistics?

- In statistics, **scenarios** can be defined as alternative versions of the same dataset.
- *For instance*, we might have a dataset with 1 observation per patient, and data on age, smoking and lung disease.
- And we might compute an alternative version of the same dataset, with the same age distribution, but all patients lifelong non–smokers.
- Lung disease outcome distributions in these two versions can be **compared** using a **statistic**, such as a mean difference, a mean ratio, a ratio between means, a median difference, or a difference between medians.
- And all of these comparisons can be presented with confidence limits and *P*–values.
- We can therefore estimate the consequences of a proposed intervention, *assuming* that an association between smoking and lung disease is **causal**, at least within each age group.

What are scenario comparison statistics?

- In statistics, **scenarios** can be defined as alternative versions of the same dataset.
- *For instance*, we might have a dataset with 1 observation per patient, and data on age, smoking and lung disease.
- And we might compute an alternative version of the same dataset, with the same age distribution, but all patients lifelong non–smokers.
- Lung disease outcome distributions in these two versions can be **compared** using a **statistic**, such as a mean difference, a mean ratio, a ratio between means, a median difference, or a difference between medians.
- And all of these comparisons can be presented with confidence limits and P–values.
- We can therefore estimate the consequences of a proposed intervention, *assuming* that an association between smoking and lung disease is **causal**, at least within each age group.

What are scenario comparison statistics?

- In statistics, **scenarios** can be defined as alternative versions of the same dataset.
- *For instance*, we might have a dataset with 1 observation per patient, and data on age, smoking and lung disease.
- And we might compute an alternative version of the same dataset, with the same age distribution, but all patients lifelong non–smokers.
- Lung disease outcome distributions in these two versions can be **compared** using a **statistic**, such as a mean difference, a mean ratio, a ratio between means, a median difference, or a difference between medians.
- And all of these comparisons can be presented with confidence limits and *P*–values.
- We can therefore estimate the consequences of a proposed intervention, *assuming* that an association between smoking and lung disease is **causal**, at least within each age group.

Example: Asthma and prenatal paracetamol in the ALSPAC cohort

- The Avon Longitudinal Study of Parents and Children (ALSPAC) is a birth cohort study, based at Bristol University, with 14060 subjects, born in the early 1990s.
- The following example revisits a subset of the dataset discussed in Shaheen *et al.*, 2010[5].
- The **outcome** of primary interest, in this example, is doctor–diagnosed asthma at any time from 0–91 months of age.
- The **exposure** is self–reported maternal paracetamol consumption during weeks 20–32 of pregnancy ("Never", "Sometimes" or "Most days/daily").
- The proposed **intervention** is presumably the replacement of paracetamol with some unspecified other drug as the analgesic of choice for pregnant mothers.
- The **candidate confounder list** was . . .

Example: Asthma and prenatal paracetamol in the ALSPAC cohort

- The Avon Longitudinal Study of Parents and Children (ALSPAC) is a birth cohort study, based at Bristol University, with 14060 subjects, born in the early 1990s.
- The following example revisits a subset of the dataset discussed in Shaheen *et al.*, 2010[5].
- The **outcome** of primary interest, in this example, is doctor–diagnosed asthma at any time from 0–91 months of age.
- The **exposure** is self–reported maternal paracetamol consumption during weeks 20–32 of pregnancy ("Never", "Sometimes" or "Most days/daily").
- The proposed **intervention** is presumably the replacement of paracetamol with some unspecified other drug as the analgesic of choice for pregnant mothers.
- The **candidate confounder list** was . . .

Example: Asthma and prenatal paracetamol in the ALSPAC cohort

- The Avon Longitudinal Study of Parents and Children (ALSPAC) is a birth cohort study, based at Bristol University, with 14060 subjects, born in the early 1990s.
- The following example revisits a subset of the dataset discussed in Shaheen *et al.*, 2010[5].
- The **outcome** of primary interest, in this example, is doctor–diagnosed asthma at any time from 0–91 months of age.
- The **exposure** is self–reported maternal paracetamol consumption during weeks 20–32 of pregnancy ("Never", "Sometimes" or "Most days/daily").
- The proposed **intervention** is presumably the replacement of paracetamol with some unspecified other drug as the analgesic of choice for pregnant mothers.
- The **candidate confounder list** was ...

Example: Asthma and prenatal paracetamol in the ALSPAC cohort

- The Avon Longitudinal Study of Parents and Children (ALSPAC) is a birth cohort study, based at Bristol University, with 14060 subjects, born in the early 1990s.
- The following example revisits a subset of the dataset discussed in Shaheen *et al.*, 2010[5].
- The **outcome** of primary interest, in this example, is doctor–diagnosed asthma at any time from 0–91 months of age.
- The **exposure** is self–reported maternal paracetamol consumption during weeks 20–32 of pregnancy ("Never", "Sometimes" or "Most days/daily").
- The proposed **intervention** is presumably the replacement of paracetamol with some unspecified other drug as the analgesic of choice for pregnant mothers.
- The **candidate confounder list** was …

Example: Asthma and prenatal paracetamol in the ALSPAC cohort

- The Avon Longitudinal Study of Parents and Children (ALSPAC) is a birth cohort study, based at Bristol University, with 14060 subjects, born in the early 1990s.
- The following example revisits a subset of the dataset discussed in Shaheen *et al.*, 2010[5].
- The **outcome** of primary interest, in this example, is doctor–diagnosed asthma at any time from 0–91 months of age.
- The **exposure** is self–reported maternal paracetamol consumption during weeks 20–32 of pregnancy ("Never", "Sometimes" or "Most days/daily").
- The proposed **intervention** is presumably the replacement of paracetamol with some unspecified other drug as the analgesic of choice for pregnant mothers.
- The **candidate confounder list** was ...

Example: Asthma and prenatal paracetamol in the ALSPAC cohort

- The Avon Longitudinal Study of Parents and Children (ALSPAC) is a birth cohort study, based at Bristol University, with 14060 subjects, born in the early 1990s.
- The following example revisits a subset of the dataset discussed in Shaheen *et al.*, 2010[5].
- The **outcome** of primary interest, in this example, is doctor–diagnosed asthma at any time from 0–91 months of age.
- The **exposure** is self–reported maternal paracetamol consumption during weeks 20–32 of pregnancy (“Never”, “Sometimes” or “Most days/daily”).
- The proposed **intervention** is presumably the replacement of paracetamol with some unspecified other drug as the analgesic of choice for pregnant mothers.
- The **candidate confounder list** was . . .

Example: Asthma and prenatal paracetamol in the ALSPAC cohort

- The Avon Longitudinal Study of Parents and Children (ALSPAC) is a birth cohort study, based at Bristol University, with 14060 subjects, born in the early 1990s.
- The following example revisits a subset of the dataset discussed in Shaheen *et al.*, 2010[5].
- The **outcome** of primary interest, in this example, is doctor–diagnosed asthma at any time from 0–91 months of age.
- The **exposure** is self–reported maternal paracetamol consumption during weeks 20–32 of pregnancy (“Never”, “Sometimes” or “Most days/daily”).
- The proposed **intervention** is presumably the replacement of paracetamol with some unspecified other drug as the analgesic of choice for pregnant mothers.
- The **candidate confounder list** was . . .

Candidate confounders

- … gender, maternal age group, prenatal tobacco exposure, maternal education, maternal housing tenure, parity, maternal anxiety group, maternal ethnic origin, multiple pregnancy, birth weight, gestational age at birth, head circumference at birth, maternal antibiotic use in pregnancy, alcohol exposure (0–8 weeks gestation, 0–18 weeks gestation, 18–32 weeks gestation, last 2 months gestation), maternal pre–pregnancy disease history (asthma, eczema, rhinoconjunctivitis, migraine), maternal infection history during pregnancy (colds/flu, urinary, other), younger siblings at 7 years, pets in first year, breast feeding in first 6 months, day care in first year, damp in home, weekend environmental tobacco exposure in first year, child's BMI at 7 years.
- Most of these are not likely to be "causally upstream" from prenatal paracetamol or asthma, but might indicate aspects of health and/or wealth that might influence both, and which are probably unaffected by analgesic substitution.

Candidate confounders

- … gender, maternal age group, prenatal tobacco exposure, maternal education, maternal housing tenure, parity, maternal anxiety group, maternal ethnic origin, multiple pregnancy, birth weight, gestational age at birth, head circumference at birth, maternal antibiotic use in pregnancy, alcohol exposure (0–8 weeks gestation, 0–18 weeks gestation, 18–32 weeks gestation, last 2 months gestation), maternal pre–pregnancy disease history (asthma, eczema, rhinoconjunctivitis, migraine), maternal infection history during pregnancy (colds/flu, urinary, other), younger siblings at 7 years, pets in first year, breast feeding in first 6 months, day care in first year, damp in home, weekend environmental tobacco exposure in first year, child's BMI at 7 years.
- Most of these are not likely to be "causally upstream" from prenatal paracetamol or asthma, but might indicate aspects of health and/or wealth that might influence both, and which are probably unaffected by analgesic substitution.

Candidate confounders

- … gender, maternal age group, prenatal tobacco exposure, maternal education, maternal housing tenure, parity, maternal anxiety group, maternal ethnic origin, multiple pregnancy, birth weight, gestational age at birth, head circumference at birth, maternal antibiotic use in pregnancy, alcohol exposure (0–8 weeks gestation, 0–18 weeks gestation, 18–32 weeks gestation, last 2 months gestation), maternal pre–pregnancy disease history (asthma, eczema, rhinoconjunctivitis, migraine), maternal infection history during pregnancy (colds/flu, urinary, other), younger siblings at 7 years, pets in first year, breast feeding in first 6 months, day care in first year, damp in home, weekend environmental tobacco exposure in first year, child's BMI at 7 years.
- Most of these are not likely to be "causally upstream" from prenatal paracetamol or asthma, but might indicate aspects of health and/or wealth that might influence both, and which are probably unaffected by analgesic substitution.

Confounder adjustment: Propensity scores

- In the ALSPAC cohort, the mothers of 12127 children gave information on paracetamol use during weeks 20–32 of pregnancy.
- We defined a paracetamol–propensity score for each of these children, based on the listed confounders, using an ordinal logistic regression model, as proposed by Lu *et al.*, 2001[3].
- These 12127 children were grouped into 32 nearly–equal propensity groups, based on the propensity score.
- All of these propensity groups were represented in the subset of 7704 of these children with data on doctor–diagnosed asthma at ages up to 91 months.

Confounder adjustment: Propensity scores

- In the ALSPAC cohort, the mothers of 12127 children gave information on paracetamol use during weeks 20–32 of pregnancy.
- We defined a paracetamol–propensity score for each of these children, based on the listed confounders, using an ordinal logistic regression model, as proposed by Lu *et al.*, 2001[3].
- These 12127 children were grouped into 32 nearly–equal propensity groups, based on the propensity score.
- All of these propensity groups were represented in the subset of 7704 of these children with data on doctor–diagnosed asthma at ages up to 91 months.

Confounder adjustment: Propensity scores

- In the ALSPAC cohort, the mothers of 12127 children gave information on paracetamol use during weeks 20–32 of pregnancy.
- We defined a paracetamol–propensity score for each of these children, based on the listed confounders, using an ordinal logistic regression model, as proposed by Lu *et al.*, 2001[3].
- These 12127 children were grouped into 32 nearly–equal propensity groups, based on the propensity score.
- All of these propensity groups were represented in the subset of 7704 of these children with data on doctor–diagnosed asthma at ages up to 91 months.

Confounder adjustment: Propensity scores

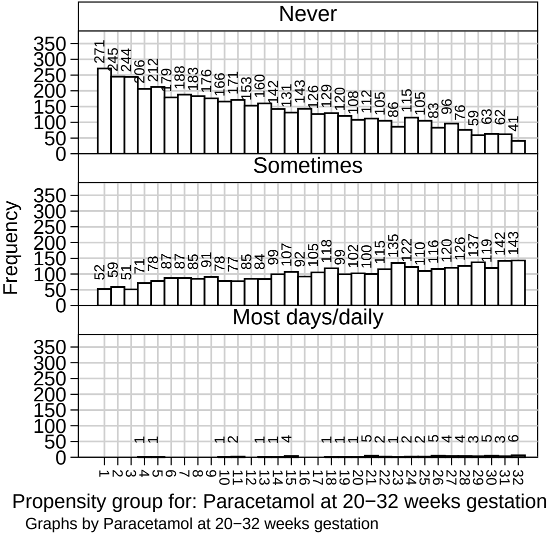
- In the ALSPAC cohort, the mothers of 12127 children gave information on paracetamol use during weeks 20–32 of pregnancy.
- We defined a paracetamol–propensity score for each of these children, based on the listed confounders, using an ordinal logistic regression model, as proposed by Lu *et al.*, 2001[3].
- These 12127 children were grouped into 32 nearly–equal propensity groups, based on the propensity score.
- All of these propensity groups were represented in the subset of 7704 of these children with data on doctor–diagnosed asthma at ages up to 91 months.

Confounder adjustment: Propensity scores

- In the ALSPAC cohort, the mothers of 12127 children gave information on paracetamol use during weeks 20–32 of pregnancy.
- We defined a paracetamol–propensity score for each of these children, based on the listed confounders, using an ordinal logistic regression model, as proposed by Lu *et al.*, 2001[3].
- These 12127 children were grouped into 32 nearly–equal propensity groups, based on the propensity score.
- All of these propensity groups were represented in the subset of 7704 of these children with data on doctor–diagnosed asthma at ages up to 91 months.

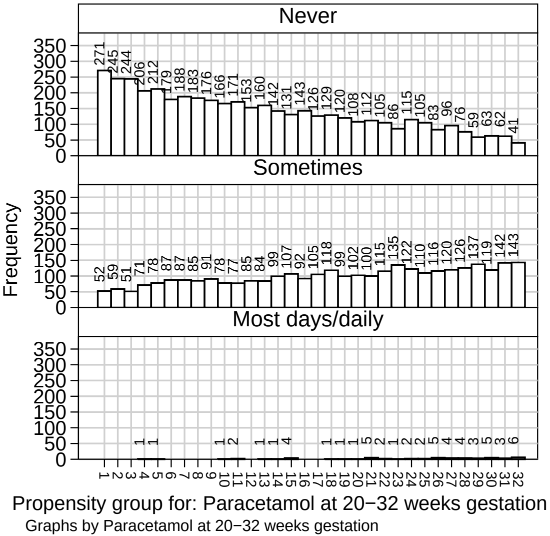
Histograms of paracetamol propensity group by paracetamol exposure

- Paracetamol propensity seems to predict paracetamol exposure, but not *too* well.
- All 32 propensity groups are represented in the unexposed group.
- *Therefore*, we should be able to contrast asthma risk between the real world and a fantasy scenario, with the same propensity distribution but no exposure.



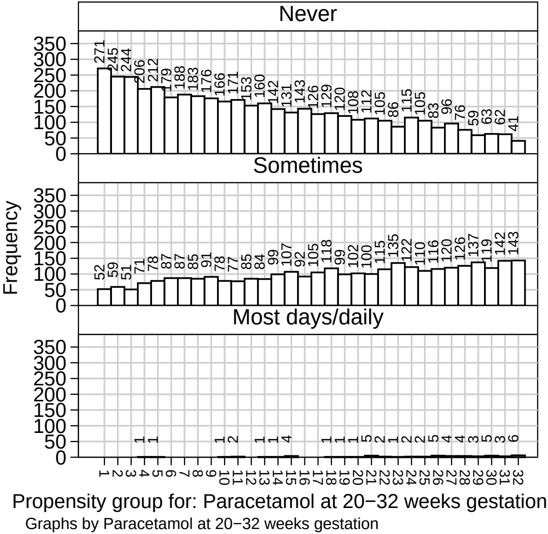
Histograms of paracetamol propensity group by paracetamol exposure

- Paracetamol propensity seems to predict paracetamol exposure, but not *too* well.
- All 32 propensity groups are represented in the unexposed group.
- *Therefore*, we should be able to contrast asthma risk between the real world and a fantasy scenario, with the same propensity distribution but no exposure.



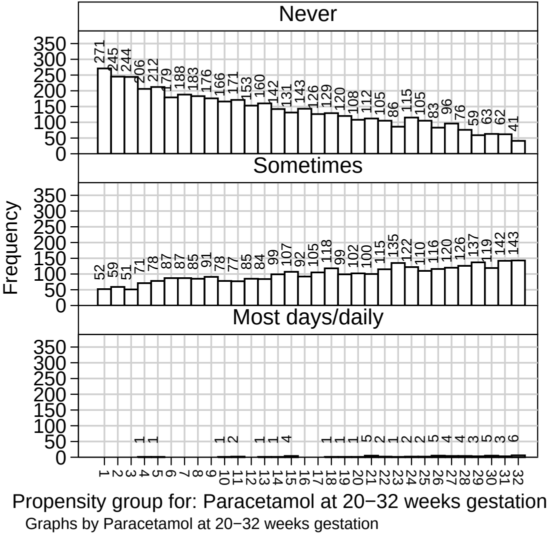
Histograms of paracetamol propensity group by paracetamol exposure

- Paracetamol propensity seems to predict paracetamol exposure, but not *too* well.
- All 32 propensity groups are represented in the unexposed group.
- *Therefore*, we should be able to contrast asthma risk between the real world and a fantasy scenario, with the same propensity distribution but no exposure.



Histograms of paracetamol propensity group by paracetamol exposure

- Paracetamol propensity seems to predict paracetamol exposure, but not *too* well.
- All 32 propensity groups are represented in the unexposed group.
- *Therefore*, we should be able to contrast asthma risk between the real world and a fantasy scenario, with the same propensity distribution but no exposure.



Method 1: Logistic regression and the population attributable fraction (Greenland and Drescher, 1993[2])

- We fitted a propensity–adjusted logistic regresssion model, with a baseline odds of asthma for each propensity group, and an odds ratio for each non–zero paracetamol exposure level.
- This model does not assume that paracetamol effects are "linear" per exposure category, but assumes that they are the same in all paracetamol–propensity groups.
- We then used the `punaf` add–on package, available in Version 11 of Stata, to estimate the **population unattributable fraction (PUF),** assuming this model.
- The PUF is a ratio between asthma risks in a fantasy scenario, with the same propensity distribution and no exposure, and asthma risks in the real world.
- The PUF, and its confidence limits, are subtracted from 1 to give a confidence interval for the **population attributable fraction (PAF)**, assuming the same model.

Method 1: Logistic regression and the population attributable fraction (Greenland and Drescher, 1993[2])

- We fitted a propensity–adjusted logistic regresssion model, with a baseline odds of asthma for each propensity group, and an odds ratio for each non–zero paracetamol exposure level.
- This model does not assume that paracetamol effects are "linear" per exposure category, but assumes that they are the same in all paracetamol–propensity groups.
- We then used the `punaf` add–on package, available in Version 11 of Stata, to estimate the **population unattributable fraction (PUF),** assuming this model.
- The PUF is a ratio between asthma risks in a fantasy scenario, with the same propensity distribution and no exposure, and asthma risks in the real world.
- The PUF, and its confidence limits, are subtracted from 1 to give a confidence interval for the **population attributable fraction (PAF)**, assuming the same model.

Method 1: Logistic regression and the population attributable fraction (Greenland and Drescher, 1993[2])

- We fitted a propensity–adjusted logistic regresssion model, with a baseline odds of asthma for each propensity group, and an odds ratio for each non–zero paracetamol exposure level.
- This model does not assume that paracetamol effects are "linear" per exposure category, but assumes that they are the same in all paracetamol–propensity groups.
- We then used the `punaf` add–on package, available in Version 11 of Stata, to estimate the **population unattributable fraction (PUF),** assuming this model.
- The PUF is a ratio between asthma risks in a fantasy scenario, with the same propensity distribution and no exposure, and asthma risks in the real world.
- The PUF, and its confidence limits, are subtracted from 1 to give a confidence interval for the **population attributable fraction (PAF)**, assuming the same model.

Method 1: Logistic regression and the population attributable fraction (Greenland and Drescher, 1993[2])

- We fitted a propensity–adjusted logistic regresssion model, with a baseline odds of asthma for each propensity group, and an odds ratio for each non–zero paracetamol exposure level.
- This model does not assume that paracetamol effects are "linear" per exposure category, but assumes that they are the same in all paracetamol–propensity groups.
- We then used the `punaf` add–on package, available in Version 11 of Stata, to estimate the **population unattributable fraction (PUF),** assuming this model.
- The PUF is a ratio between asthma risks in a fantasy scenario, with the same propensity distribution and no exposure, and asthma risks in the real world.
- The PUF, and its confidence limits, are subtracted from 1 to give a confidence interval for the **population attributable fraction (PAF)**, assuming the same model.

Method 1: Logistic regression and the population attributable fraction (Greenland and Drescher, 1993[2])

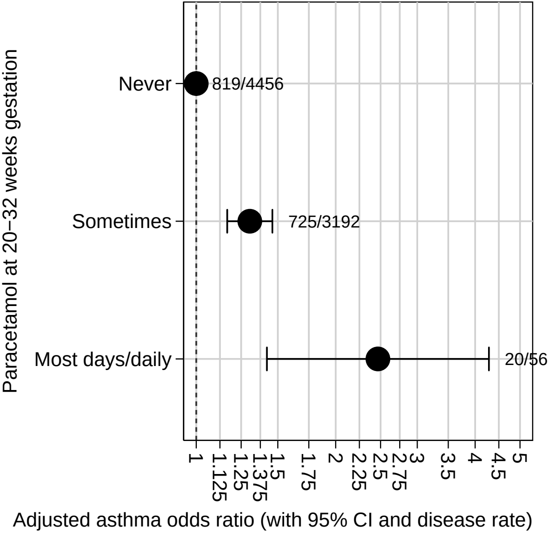
- We fitted a propensity–adjusted logistic regresssion model, with a baseline odds of asthma for each propensity group, and an odds ratio for each non–zero paracetamol exposure level.
- This model does not assume that paracetamol effects are "linear" per exposure category, but assumes that they are the same in all paracetamol–propensity groups.
- We then used the `punaf` add–on package, available in Version 11 of Stata, to estimate the **population unattributable fraction (PUF),** assuming this model.
- The PUF is a ratio between asthma risks in a fantasy scenario, with the same propensity distribution and no exposure, and asthma risks in the real world.
- The PUF, and its confidence limits, are subtracted from 1 to give a confidence interval for the **population attributable fraction (PAF)**, assuming the same model.

Method 1: Logistic regression and the population attributable fraction (Greenland and Drescher, 1993[2])

- We fitted a propensity–adjusted logistic regresssion model, with a baseline odds of asthma for each propensity group, and an odds ratio for each non–zero paracetamol exposure level.
- This model does not assume that paracetamol effects are “linear” per exposure category, but assumes that they are the same in all paracetamol–propensity groups.
- We then used the `punaf` add–on package, available in Version 11 of Stata, to estimate the **population unattributable fraction (PUF),** assuming this model.
- The PUF is a ratio between asthma risks in a fantasy scenario, with the same propensity distribution and no exposure, and asthma risks in the real world.
- The PUF, and its confidence limits, are subtracted from 1 to give a confidence interval for the **population attributable fraction (PAF)**, assuming the same model.

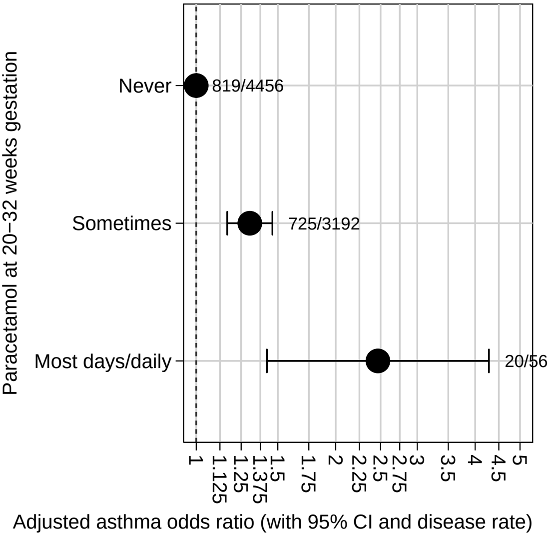
Adjusted odds ratios of asthma with respect to paracetamol exposure

- Paracetamol exposure is associated with increased asthma, even allowing for paracetamol propensity.
- *However*, we would like to have a summary of overall trend.
- And it would be even better if it indicated how much good might be done by replacing paracetamol with a "harmless" analgesic.



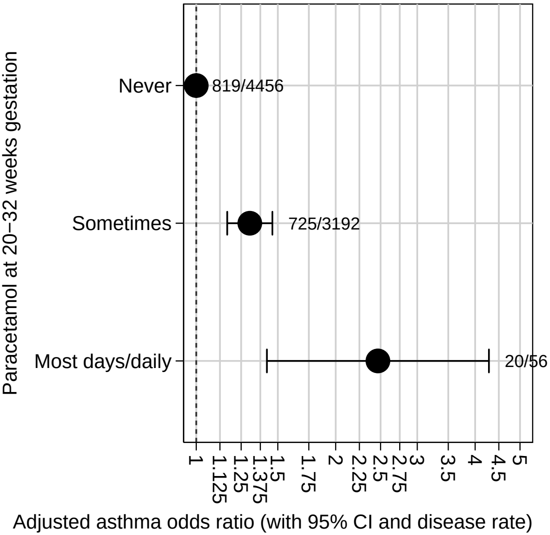
Adjusted odds ratios of asthma with respect to paracetamol exposure

- Paracetamol exposure is associated with increased asthma, even allowing for paracetamol propensity.
- *However*, we would like to have a summary of overall trend.
- And it would be even better if it indicated how much good might be done by replacing paracetamol with a “harmless” analgesic.



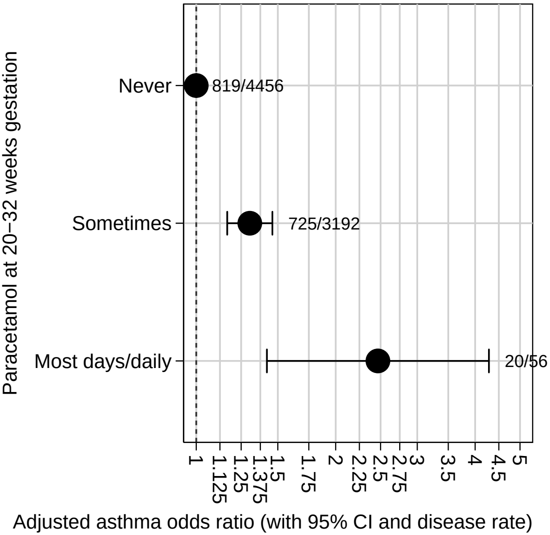
Adjusted odds ratios of asthma with respect to paracetamol exposure

- Paracetamol exposure is associated with increased asthma, even allowing for paracetamol propensity.
- *However*, we would like to have a summary of overall trend.
- And it would be even better if it indicated how much good might be done by replacing paracetamol with a “harmless” analgesic.



Adjusted odds ratios of asthma with respect to paracetamol exposure

- Paracetamol exposure is associated with increased asthma, even allowing for paracetamol propensity.
- *However*, we would like to have a summary of overall trend.
- And it would be even better if it indicated how much good might be done by replacing paracetamol with a "harmless" analgesic.



Scenario means and population unattributable and attributable fractions

After fitting the logistic regression model, we use `punaf` to compute the scenario means (asthma risks) and population unattributable and attributable fractions:

```
. punaf, atspec(para32g=0) eform post;

Confidence intervals for the scenario means
under Scenario 0 (baseline) and Scenario 1 (specified by atspec() option)
and for the population unattributable faction (PUF)
Total number of observations used: 7704
------------------------------------------------------------------------------
             | Mean/Ratio   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
  Scenario_0 |   .2030114    .004558   -71.02   0.000     .1942717    .2121444
  Scenario_1 |   .1889629   .0060648   -51.91   0.000     .1774423    .2012314
         PUF |   .9307992    .020928    -3.19   0.001     .8906717    .9727347
------------------------------------------------------------------------------

95% CI for the population attributable fraction (PAF)
              Estimate     Minimum     Maximum
         PAF  .06920077  .02726535  .10932832
```

“Scenario_0” is the real world of our sample. “Scenario_1” is the fantasy sample with the same propensity distribution and no exposure. And the PUF is the “Scenario_1”/“Scenario_0” asthma risk ratio, which can be subtracted from 1 to derive the PAF.

Method 1: Summary of results

- Under "Scenario 0" (the real world), 20.30% of subjects (95% CI, 19.43% to 21.21%) had ever had doctor–diagnosed asthma by the age of 91 months.
- Under "Scenario 1" (the same distribution of "paracetamol proneness" but no paracetamol exposure), the risk might be 18.90% (95% CI, 17.74% to 20.12%).
- The ratio of the second risk to the first (the PUF) is .9308 (95% CI, .8907 to .9727), indicating that 89 to 97 percent of the risk would remain if paracetamol exposure was eliminated but confounding factors stayed the same.
- *So* the PAF (the fraction of lifetime asthma risk attributable to prenatal paracetamol exposure) is 6.92% of the total risk (95% CI, 2.73% to 10.93%).
- Most people would *probably* understand this message more easily than the odds ratios.

Method 1: Summary of results

- Under "Scenario 0" (the real world), 20.30% of subjects (95% CI, 19.43% to 21.21%) had ever had doctor–diagnosed asthma by the age of 91 months.
- Under "Scenario 1" (the same distribution of "paracetamol proneness" but no paracetamol exposure), the risk might be 18.90% (95% CI, 17.74% to 20.12%).
- The ratio of the second risk to the first (the PUF) is .9308 (95% CI, .8907 to .9727), indicating that 89 to 97 percent of the risk would remain if paracetamol exposure was eliminated but confounding factors stayed the same.
- *So* the PAF (the fraction of lifetime asthma risk attributable to prenatal paracetamol exposure) is 6.92% of the total risk (95% CI, 2.73% to 10.93%).
- Most people would *probably* understand this message more easily than the odds ratios.

Method 1: Summary of results

- Under "Scenario 0" (the real world), 20.30% of subjects (95% CI, 19.43% to 21.21%) had ever had doctor–diagnosed asthma by the age of 91 months.
- Under "Scenario 1" (the same distribution of "paracetamol proneness" but no paracetamol exposure), the risk might be 18.90% (95% CI, 17.74% to 20.12%).
- The ratio of the second risk to the first (the PUF) is .9308 (95% CI, .8907 to .9727), indicating that 89 to 97 percent of the risk would remain if paracetamol exposure was eliminated but confounding factors stayed the same.
- *So* the PAF (the fraction of lifetime asthma risk attributable to prenatal paracetamol exposure) is 6.92% of the total risk (95% CI, 2.73% to 10.93%).
- Most people would *probably* understand this message more easily than the odds ratios.

Method 1: Summary of results

- Under "Scenario 0" (the real world), 20.30% of subjects (95% CI, 19.43% to 21.21%) had ever had doctor–diagnosed asthma by the age of 91 months.
- Under "Scenario 1" (the same distribution of "paracetamol proneness" but no paracetamol exposure), the risk might be 18.90% (95% CI, 17.74% to 20.12%).
- The ratio of the second risk to the first (the PUF) is .9308 (95% CI, .8907 to .9727), indicating that 89 to 97 percent of the risk would remain if paracetamol exposure was eliminated but confounding factors stayed the same.
- *So* the PAF (the fraction of lifetime asthma risk attributable to prenatal paracetamol exposure) is 6.92% of the total risk (95% CI, 2.73% to 10.93%).
- Most people would *probably* understand this message more easily than the odds ratios.

Method 1: Summary of results

- Under "Scenario 0" (the real world), 20.30% of subjects (95% CI, 19.43% to 21.21%) had ever had doctor–diagnosed asthma by the age of 91 months.
- Under "Scenario 1" (the same distribution of "paracetamol proneness" but no paracetamol exposure), the risk might be 18.90% (95% CI, 17.74% to 20.12%).
- The ratio of the second risk to the first (the PUF) is .9308 (95% CI, .8907 to .9727), indicating that 89 to 97 percent of the risk would remain if paracetamol exposure was eliminated but confounding factors stayed the same.
- *So* the PAF (the fraction of lifetime asthma risk attributable to prenatal paracetamol exposure) is 6.92% of the total risk (95% CI, 2.73% to 10.93%).
- Most people would *probably* understand this message more easily than the odds ratios.

Method 1: Summary of results

- Under "Scenario 0" (the real world), 20.30% of subjects (95% CI, 19.43% to 21.21%) had ever had doctor–diagnosed asthma by the age of 91 months.
- Under "Scenario 1" (the same distribution of "paracetamol proneness" but no paracetamol exposure), the risk might be 18.90% (95% CI, 17.74% to 20.12%).
- The ratio of the second risk to the first (the PUF) is .9308 (95% CI, .8907 to .9727), indicating that 89 to 97 percent of the risk would remain if paracetamol exposure was eliminated but confounding factors stayed the same.
- *So* the PAF (the fraction of lifetime asthma risk attributable to prenatal paracetamol exposure) is 6.92% of the total risk (95% CI, 2.73% to 10.93%).
- Most people would *probably* understand this message more easily than the odds ratios.

Method 2: Direct standardization and the population attributable risk (Newson, 2006[4])

- The **population attributable risk (PAR)** is defined by Gordis (2000)[1] as a scenario *difference* between disease risks.
- "Scenario 0" is the real world of the sample we have, and "Scenario 1" is a fantasy scenario with no exposure to the risk factor.
- This concept is easily generalized, using direct standardization, to the case where Scenarios 0 and 1 are assumed to have the same distribution of one or more categorical confounders.
- In this case, the alternative scenarios (versions of the data) are distinguished by alternative sampling–probability weights.
- These sampling–probability weights are 1 for all subjects in Scenario 0.
- In Scenario 1, they are 0 for exposed subjects, and equal, in an unexposed subject, to the ratio of its confounder–group frequency in all subjects to its confounder–group frequency in unexposed subjects.

Method 2: Direct standardization and the population attributable risk (Newson, 2006[4])

- The **population attributable risk (PAR)** is defined by Gordis (2000)[1] as a scenario *difference* between disease risks.
- "Scenario 0" is the real world of the sample we have, and "Scenario 1" is a fantasy scenario with no exposure to the risk factor.
- This concept is easily generalized, using direct standardization, to the case where Scenarios 0 and 1 are assumed to have the same distribution of one or more categorical confounders.
- In this case, the alternative scenarios (versions of the data) are distinguished by alternative sampling–probability weights.
- These sampling–probability weights are 1 for all subjects in Scenario 0.
- In Scenario 1, they are 0 for exposed subjects, and equal, in an unexposed subject, to the ratio of its confounder–group frequency in all subjects to its confounder–group frequency in unexposed subjects.

Method 2: Direct standardization and the population attributable risk (Newson, 2006[4])

- The **population attributable risk (PAR)** is defined by Gordis (2000)[1] as a scenario *difference* between disease risks.
- "Scenario 0" is the real world of the sample we have, and "Scenario 1" is a fantasy scenario with no exposure to the risk factor.
- This concept is easily generalized, using direct standardization, to the case where Scenarios 0 and 1 are assumed to have the same distribution of one or more categorical confounders.
- In this case, the alternative scenarios (versions of the data) are distinguished by alternative sampling–probability weights.
- These sampling–probability weights are 1 for all subjects in Scenario 0.
- In Scenario 1, they are 0 for exposed subjects, and equal, in an unexposed subject, to the ratio of its confounder–group frequency in all subjects to its confounder–group frequency in unexposed subjects.

Method 2: Direct standardization and the population attributable risk (Newson, 2006[4])

- The **population attributable risk (PAR)** is defined by Gordis (2000)[1] as a scenario *difference* between disease risks.
- "Scenario 0" is the real world of the sample we have, and "Scenario 1" is a fantasy scenario with no exposure to the risk factor.
- This concept is easily generalized, using direct standardization, to the case where Scenarios 0 and 1 are assumed to have the same distribution of one or more categorical confounders.
- In this case, the alternative scenarios (versions of the data) are distinguished by alternative sampling–probability weights.
- These sampling–probability weights are 1 for all subjects in Scenario 0.
- In Scenario 1, they are 0 for exposed subjects, and equal, in an unexposed subject, to the ratio of its confounder–group frequency in all subjects to its confounder–group frequency in unexposed subjects.

Method 2: Direct standardization and the population attributable risk (Newson, 2006[4])

- The **population attributable risk (PAR)** is defined by Gordis (2000)[1] as a scenario *difference* between disease risks.
- "Scenario 0" is the real world of the sample we have, and "Scenario 1" is a fantasy scenario with no exposure to the risk factor.
- This concept is easily generalized, using direct standardization, to the case where Scenarios 0 and 1 are assumed to have the same distribution of one or more categorical confounders.
- In this case, the alternative scenarios (versions of the data) are distinguished by alternative sampling–probability weights.
- These sampling–probability weights are 1 for all subjects in Scenario 0.
- In Scenario 1, they are 0 for exposed subjects, and equal, in an unexposed subject, to the ratio of its confounder–group frequency in all subjects to its confounder–group frequency in unexposed subjects.

Method 2: Direct standardization and the population attributable risk (Newson, 2006[4])

- The **population attributable risk (PAR)** is defined by Gordis (2000)[1] as a scenario *difference* between disease risks.
- "Scenario 0" is the real world of the sample we have, and "Scenario 1" is a fantasy scenario with no exposure to the risk factor.
- This concept is easily generalized, using direct standardization, to the case where Scenarios 0 and 1 are assumed to have the same distribution of one or more categorical confounders.
- In this case, the alternative scenarios (versions of the data) are distinguished by alternative sampling–probability weights.
- These sampling–probability weights are 1 for all subjects in Scenario 0.
- In Scenario 1, they are 0 for exposed subjects, and equal, in an unexposed subject, to the ratio of its confounder–group frequency in all subjects to its confounder–group frequency in unexposed subjects.

Method 2: Direct standardization and the population attributable risk (Newson, 2006[4])

- The **population attributable risk (PAR)** is defined by Gordis (2000)[1] as a scenario *difference* between disease risks.
- "Scenario 0" is the real world of the sample we have, and "Scenario 1" is a fantasy scenario with no exposure to the risk factor.
- This concept is easily generalized, using direct standardization, to the case where Scenarios 0 and 1 are assumed to have the same distribution of one or more categorical confounders.
- In this case, the alternative scenarios (versions of the data) are distinguished by alternative sampling–probability weights.
- These sampling–probability weights are 1 for all subjects in Scenario 0.
- In Scenario 1, they are 0 for exposed subjects, and equal, in an unexposed subject, to the ratio of its confounder–group frequency in all subjects to its confounder–group frequency in unexposed subjects.

The population attributable risk as a rank statistic

- Arguably, a difference between proportions is more naturally a rank statistic than a regression statistic, although it can be either.
- It is defined as a special case of Somers' D, which in general is a difference between probabilities of concordance and discordance.
- Sensible Normalizing and variance–stabilizing transforms for a difference between proportions include the arcsine and the hyperbolic arctangent (also known as Fisher's z).
- (This is in contrast to the proportions themselves, for which a sensible transformation is the log odds.)
- The `scsomersd` add–on package, available in Version 10 of Stata, estimates a range of scenario–comparison rank statistics, including the PAR, using scenario–specific sampling–probability weights.

The population attributable risk as a rank statistic

- Arguably, a difference between proportions is more naturally a rank statistic than a regression statistic, although it can be either.
- It is defined as a special case of Somers' D, which in general is a difference between probabilities of concordance and discordance.
- Sensible Normalizing and variance–stabilizing transforms for a difference between proportions include the arcsine and the hyperbolic arctangent (also known as Fisher's z).
- (This is in contrast to the proportions themselves, for which a sensible transformation is the log odds.)
- The `scsomersd` add–on package, available in Version 10 of Stata, estimates a range of scenario–comparison rank statistics, including the PAR, using scenario–specific sampling–probability weights.

The population attributable risk as a rank statistic

- Arguably, a difference between proportions is more naturally a rank statistic than a regression statistic, although it can be either.
- It is defined as a special case of Somers' D, which in general is a difference between probabilities of concordance and discordance.
- Sensible Normalizing and variance–stabilizing transforms for a difference between proportions include the arcsine and the hyperbolic arctangent (also known as Fisher's z).
- (This is in contrast to the proportions themselves, for which a sensible transformation is the log odds.)
- The `scsomersd` add–on package, available in Version 10 of Stata, estimates a range of scenario–comparison rank statistics, including the PAR, using scenario–specific sampling–probability weights.

The population attributable risk as a rank statistic

- Arguably, a difference between proportions is more naturally a rank statistic than a regression statistic, although it can be either.
- It is defined as a special case of Somers' D, which in general is a difference between probabilities of concordance and discordance.
- Sensible Normalizing and variance–stabilizing transforms for a difference between proportions include the arcsine and the hyperbolic arctangent (also known as Fisher's z).
- (This is in contrast to the proportions themselves, for which a sensible transformation is the log odds.)
- The `scsomersd` add–on package, available in Version 10 of Stata, estimates a range of scenario–comparison rank statistics, including the PAR, using scenario–specific sampling–probability weights.

The population attributable risk as a rank statistic

- Arguably, a difference between proportions is more naturally a rank statistic than a regression statistic, although it can be either.
- It is defined as a special case of Somers' D, which in general is a difference between probabilities of concordance and discordance.
- Sensible Normalizing and variance–stabilizing transforms for a difference between proportions include the arcsine and the hyperbolic arctangent (also known as Fisher's z).
- (This is in contrast to the proportions themselves, for which a sensible transformation is the log odds.)
- The `scsomersd` add–on package, available in Version 10 of Stata, estimates a range of scenario–comparison rank statistics, including the PAR, using scenario–specific sampling–probability weights.

The population attributable risk as a rank statistic

- Arguably, a difference between proportions is more naturally a rank statistic than a regression statistic, although it can be either.
- It is defined as a special case of Somers' D, which in general is a difference between probabilities of concordance and discordance.
- Sensible Normalizing and variance–stabilizing transforms for a difference between proportions include the arcsine and the hyperbolic arctangent (also known as Fisher's z).
- (This is in contrast to the proportions themselves, for which a sensible transformation is the log odds.)
- The `scsomersd` add–on package, available in Version 10 of Stata, estimates a range of scenario–comparison rank statistics, including the PAR, using scenario–specific sampling–probability weights.

Example: Paracetamol exposure and asthma in ALSPAC

- The exposure is late pre–natal paracetamol exposure (at any level).
- The outcome is doctor–diagnosed asthma (ever by age 91 months).
- The confounder groups are the 32 paracetamol–propensity groups.
- The sampling–probability weights in Scenario 1 (elimination of paracetamol exposure) are equal, for each unexposed subject, to the ratio between the frequencies of its propensity group in all subjects and in unexposed subjects.
- We computed the scenario risks from scenario odds, using logistic regression with sampling–probability weights, and the directly–standardized difference between these risks, using `scsomersd`.

Example: Paracetamol exposure and asthma in ALSPAC

- The exposure is late pre–natal paracetamol exposure (at any level).
- The outcome is doctor–diagnosed asthma (ever by age 91 months).
- The confounder groups are the 32 paracetamol–propensity groups.
- The sampling–probability weights in Scenario 1 (elimination of paracetamol exposure) are equal, for each unexposed subject, to the ratio between the frequencies of its propensity group in all subjects and in unexposed subjects.
- We computed the scenario risks from scenario odds, using logistic regression with sampling–probability weights, and the directly–standardized difference between these risks, using `scsomersd`.

Example: Paracetamol exposure and asthma in ALSPAC

- The exposure is late pre–natal paracetamol exposure (at any level).
- The outcome is doctor–diagnosed asthma (ever by age 91 months).
- The confounder groups are the 32 paracetamol–propensity groups.
- The sampling–probability weights in Scenario 1 (elimination of paracetamol exposure) are equal, for each unexposed subject, to the ratio between the frequencies of its propensity group in all subjects and in unexposed subjects.
- We computed the scenario risks from scenario odds, using logistic regression with sampling–probability weights, and the directly–standardized difference between these risks, using `scsomersd`.

Example: Paracetamol exposure and asthma in ALSPAC

- The exposure is late pre–natal paracetamol exposure (at any level).
- The outcome is doctor–diagnosed asthma (ever by age 91 months).
- The confounder groups are the 32 paracetamol–propensity groups.
- The sampling–probability weights in Scenario 1 (elimination of paracetamol exposure) are equal, for each unexposed subject, to the ratio between the frequencies of its propensity group in all subjects and in unexposed subjects.
- We computed the scenario risks from scenario odds, using logistic regression with sampling–probability weights, and the directly–standardized difference between these risks, using `scsomersd`.

Example: Paracetamol exposure and asthma in ALSPAC

- The exposure is late pre–natal paracetamol exposure (at any level).
- The outcome is doctor–diagnosed asthma (ever by age 91 months).
- The confounder groups are the 32 paracetamol–propensity groups.
- The sampling–probability weights in Scenario 1 (elimination of paracetamol exposure) are equal, for each unexposed subject, to the ratio between the frequencies of its propensity group in all subjects and in unexposed subjects.
- We computed the scenario risks from scenario odds, using logistic regression with sampling–probability weights, and the directly–standardized difference between these risks, using `scsomersd`.

Example: Paracetamol exposure and asthma in ALSPAC

- The exposure is late pre–natal paracetamol exposure (at any level).
- The outcome is doctor–diagnosed asthma (ever by age 91 months).
- The confounder groups are the 32 paracetamol–propensity groups.
- The sampling–probability weights in Scenario 1 (elimination of paracetamol exposure) are equal, for each unexposed subject, to the ratio between the frequencies of its propensity group in all subjects and in unexposed subjects.
- We computed the scenario risks from scenario odds, using logistic regression with sampling–probability weights, and the directly–standardized difference between these risks, using `scsomersd`.

Estimation of population attributable risk using **scsomersd**

The specification `pwei=1` specifies sampling–probability weights in Scenario 0. The option `sweight(dsweight*(exposed==0))` specifies sampling–probability weights in Scenario 1.

```
. scsomersd ddasth91 [pwei=1], sweight(dsweight*(exposed==0)) transf(z) tdist;
Von Mises Somers' D with variable: _scen0
Transformation: Fisher's z
Valid observations: 12160
Number of clusters: 7704
Degrees of freedom: 7703

Symmetric 95% CI for transformed Somers' D
                                (Std. Err. adjusted for 7704 clusters in _obs)
------------------------------------------------------------------------------
             |              Jackknife
      _scen0 |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
       _yvar |   .0114349   .0045842     2.49   0.013     .0024486    .0204213
------------------------------------------------------------------------------

Asymmetric 95% CI for untransformed Somers' D
               Somers_D     Minimum     Maximum
       _yvar   .01143442   .00244855   .02041844
```

The output is in alien–looking language. *However*, the bottom line (giving the PAR) is the part that we can communicate to pregnant mothers and health professionals.

Method 2: Scenario risks and the population attributable risk

Scenario	*N*	*Risk*	*(95%*	*CI)*	*P*
Scenario 0	7704	0.2030	(0.1942,	0.2121)	
Scenario 1	4456	0.1916	(0.1795,	0.2043)	
PAR	7704	0.0113	(0.0024,	0.0200)	.013

- We see that, *if* we substituted a "harmless" analgesic for paracetamol during pregnancy, *and* "paracetamol proneness" remained the same, *then* we might save 1% of children from ever having asthma, at least up to 91 months of age.
- This percent (and its confidence limits) will look more exciting if multiplied by the total population of children in the UK.
- Note that the asthma risk under Scenario 1 by Method 2 is slightly different from the asthma risk under Scenario 1 by Method 1, as we are no longer assuming paracetamol odds ratios to be constant between propensity groups.

Method 2: Scenario risks and the population attributable risk

Scenario	*N*	*Risk*	*(95%*	*CI)*	*P*
Scenario 0	7704	0.2030	(0.1942,	0.2121)	
Scenario 1	4456	0.1916	(0.1795,	0.2043)	
PAR	7704	0.0113	(0.0024,	0.0200)	.013

- We see that, *if* we substituted a "harmless" analgesic for paracetamol during pregnancy, *and* "paracetamol proneness" remained the same, *then* we might save 1% of children from ever having asthma, at least up to 91 months of age.
- This percent (and its confidence limits) will look more exciting if multiplied by the total population of children in the UK.
- Note that the asthma risk under Scenario 1 by Method 2 is slightly different from the asthma risk under Scenario 1 by Method 1, as we are no longer assuming paracetamol odds ratios to be constant between propensity groups.

Method 2: Scenario risks and the population attributable risk

Scenario	*N*	*Risk*	*(95%*	*CI)*	*P*
Scenario 0	7704	0.2030	(0.1942,	0.2121)	
Scenario 1	4456	0.1916	(0.1795,	0.2043)	
PAR	7704	0.0113	(0.0024,	0.0200)	.013

- We see that, *if* we substituted a "harmless" analgesic for paracetamol during pregnancy, *and* "paracetamol proneness" remained the same, *then* we might save 1% of children from ever having asthma, at least up to 91 months of age.
- This percent (and its confidence limits) will look more exciting if multiplied by the total population of children in the UK.
- Note that the asthma risk under Scenario 1 by Method 2 is slightly different from the asthma risk under Scenario 1 by Method 1, as we are no longer assuming paracetamol odds ratios to be constant between propensity groups.

Method 2: Scenario risks and the population attributable risk

Scenario	*N*	*Risk*	*(95%*	*CI)*	*P*
Scenario 0	7704	0.2030	(0.1942,	0.2121)	
Scenario 1	4456	0.1916	(0.1795,	0.2043)	
PAR	7704	0.0113	(0.0024,	0.0200)	.013

- We see that, *if* we substituted a "harmless" analgesic for paracetamol during pregnancy, *and* "paracetamol proneness" remained the same, *then* we might save 1% of children from ever having asthma, at least up to 91 months of age.
- This percent (and its confidence limits) will look more exciting if multiplied by the total population of children in the UK.
- Note that the asthma risk under Scenario 1 by Method 2 is slightly different from the asthma risk under Scenario 1 by Method 1, as we are no longer assuming paracetamol odds ratios to be constant between propensity groups.

Summary: Methods 1 and 2

- Both of these methods produce measures of overall trend that are easier to understand than odds ratios.
- Method 1 estimates scenario risks and population attributable *fractions* (proportions of *ever–asthmatics* that might have been saved).
- This is done using a logistic regression model, which assumes common paracetamol odds ratios for all sets of confounder values.
- Method 2 estimates scenario risks and population attributable *risks* (proportions of *all children* that might have been saved).
- This is done by direct standardization, and does not assume a logistic model, or any other regression model.
- *However*, direct standardization does require that unexposed subjects are represented in each paracetamol–propensity group.

Summary: Methods 1 and 2

- Both of these methods produce measures of overall trend that are easier to understand than odds ratios.
- Method 1 estimates scenario risks and population attributable *fractions* (proportions of *ever–asthmatics* that might have been saved).
- This is done using a logistic regression model, which assumes common paracetamol odds ratios for all sets of confounder values.
- Method 2 estimates scenario risks and population attributable *risks* (proportions of *all children* that might have been saved).
- This is done by direct standardization, and does not assume a logistic model, or any other regression model.
- *However*, direct standardization does require that unexposed subjects are represented in each paracetamol–propensity group.

Summary: Methods 1 and 2

- Both of these methods produce measures of overall trend that are easier to understand than odds ratios.
- Method 1 estimates scenario risks and population attributable *fractions* (proportions of *ever–asthmatics* that might have been saved).
- This is done using a logistic regression model, which assumes common paracetamol odds ratios for all sets of confounder values.
- Method 2 estimates scenario risks and population attributable *risks* (proportions of *all children* that might have been saved).
- This is done by direct standardization, and does not assume a logistic model, or any other regression model.
- *However*, direct standardization does require that unexposed subjects are represented in each paracetamol–propensity group.

Summary: Methods 1 and 2

- Both of these methods produce measures of overall trend that are easier to understand than odds ratios.
- Method 1 estimates scenario risks and population attributable *fractions* (proportions of *ever–asthmatics* that might have been saved).
- This is done using a logistic regression model, which assumes common paracetamol odds ratios for all sets of confounder values.
- Method 2 estimates scenario risks and population attributable *risks* (proportions of *all children* that might have been saved).
- This is done by direct standardization, and does not assume a logistic model, or any other regression model.
- *However*, direct standardization does require that unexposed subjects are represented in each paracetamol–propensity group.

Summary: Methods 1 and 2

- Both of these methods produce measures of overall trend that are easier to understand than odds ratios.
- Method 1 estimates scenario risks and population attributable *fractions* (proportions of *ever–asthmatics* that might have been saved).
- This is done using a logistic regression model, which assumes common paracetamol odds ratios for all sets of confounder values.
- Method 2 estimates scenario risks and population attributable *risks* (proportions of *all children* that might have been saved).
- This is done by direct standardization, and does not assume a logistic model, or any other regression model.
- *However*, direct standardization does require that unexposed subjects are represented in each paracetamol–propensity group.

Summary: Methods 1 and 2

- Both of these methods produce measures of overall trend that are easier to understand than odds ratios.
- Method 1 estimates scenario risks and population attributable *fractions* (proportions of *ever–asthmatics* that might have been saved).
- This is done using a logistic regression model, which assumes common paracetamol odds ratios for all sets of confounder values.
- Method 2 estimates scenario risks and population attributable *risks* (proportions of *all children* that might have been saved).
- This is done by direct standardization, and does not assume a logistic model, or any other regression model.
- *However*, direct standardization does require that unexposed subjects are represented in each paracetamol–propensity group.

Summary: Methods 1 and 2

- Both of these methods produce measures of overall trend that are easier to understand than odds ratios.
- Method 1 estimates scenario risks and population attributable *fractions* (proportions of *ever–asthmatics* that might have been saved).
- This is done using a logistic regression model, which assumes common paracetamol odds ratios for all sets of confounder values.
- Method 2 estimates scenario risks and population attributable *risks* (proportions of *all children* that might have been saved).
- This is done by direct standardization, and does not assume a logistic model, or any other regression model.
- *However*, direct standardization does require that unexposed subjects are represented in each paracetamol–propensity group.

Future improvements

- I plan to extend `punaf` to estimate population unattributable and attributable fractions for case–control studies.
- It would also be good to make all these methods easier for non–technical people to use.

Future improvements

- I plan to extend `punaf` to estimate population unattributable and attributable fractions for case–control studies.
- It would also be good to make all these methods easier for non–technical people to use.

Future improvements

- I plan to extend `punaf` to estimate population unattributable and attributable fractions for case–control studies.
- It would also be good to make all these methods easier for non–technical people to use.

References

[1] Gordis L. *Epidemiology. 2nd ed.* 2000; Philadelphia, PA: W. B. Saunders.

[2] Greenland S, Drescher K. Maximum likelihood estimation of the attributable fraction from logistic models. *Biometrics* 1993; **49(3)**: 865–872.

[3] Lu B, Zanutto E, Hornik R, Rosenbaum PR. Matching with doses in an observational study of a media campaign against drug abuse. *Journal of the American Statistical Association* 2001; **96**: 1245–1253.

[4] Newson R. Confidence intervals for rank statistics: Somers' *D* and extensions. *The Stata Journal* 2006; **6(3)**: 309–334.

[5] Shaheen SO, Newson RB, SM, Rose-Zerilli MJ,Holloway JW, Henderson AJ. Prenatal and infant acetaminophen exposure, antioxidant gene polymorphisms and childhood asthma. *Journal of Allergy and Clinical Immunology* 2010, in press. doi: 10.1016/j.jaci.2010.08.047.