

# Data Extraction Public Health - Overall study

Alternative to EPPI to extract data from papers for the public health sector. This form only asks for things that the whole study should have a common answer to. After this Form, please continue in the treatment arm form, filling it in once for each relevant treatment arm, and for each outcome in the treatment arms, please fill in the form concerning outcomes

When some required data is unknown, please put in 7777.

\* Required

1. ID of person extracting data \*

2. Year Of Publication \*

Number must be between 1980 ~ 2021

3. Item ID internal \*

*You can find this in EPPI (search by first author for speed)*

Number must be between 10000000 ~ 99999999

4. Publication type \*

- Peer-reviewed journal or journal
- Report/working paper/grey literature
- Can't tell

5. Origin of Intervention \*

- Community-based
- NGO
- Local/National Government
- Foreign Government
- Other

6. Is the sampling frame clearly defined? \*

*What is the defined population from which the sample is drawn? Do the authors say who could potentially be chosen as sample?*

- Yes
- No
- Can't tell

7. Does the sampling frame include at least 1000 beneficiaries or covers an administrative area larger than a village? \*

- Yes
- No
- Can't tell

8. Is the sample randomly drawn from this sampling frame? \*

- Yes
- No
- Can't tell

9. Number of Treatment arms \*

*Only count relevant treatment arms*

- 1
- 2
- 3
- 4
- 5
- 6
- 7
- 8
- 9
- 10

10. Notes

---

This content is neither created nor endorsed by Microsoft. The data you submit will be sent to the form owner.

 Microsoft Forms

# Data Extraction Public Health - Treatment Arm

Alternative to EPPI to extract data from papers for the public health sector. This form only asks for things that the treatment arm should have a common answer to. After this Form, please continue in the outcome form, filling it in once for each relevant outcome for this treatment arm. If you haven't filled in the form for the overall study data, please do that before starting with the treatment arms.

When some required data is unknown, please put in 7777

\* Required

1. ID of person extracting data \*

2. Year Of Publication \*

Number must be between 1980 ~ 2021

3. Item ID internal \*

*You can find this in EPPI (search by first author for speed)*

Number must be between 10000000 ~ 99999999

4. Which treatment arm are you coding? \*

- treatment arm 1
- treatment arm 2
- treatment arm 3
- treatment arm 4
- treatment arm 5
- treatment arm 6
- treatment arm 7
- treatment arm 8
- treatment arm 9
- treatment arm 10

5. Description of Intervention(s) of this treatment arm \*

6. Type of Interventions Part of Treatment Arm

- coercion
- education
- enablement
- incentives
- modelling
- persuasion
- physical (environment) restructuring
- restriction
- social (environment) restructuring
- training

7. Timing of the implementation of the intervention \*

- One time
- 1 to 12 months
- 1 to 3 years
- More than 3 years
- Can't tell

8. Blinding of the trial (if applicable)

- Blinding of Participants
- Blinding of Clinicians
- Blinding of Data Collectors
- Blinding of Data Analysts
- Can't tell

9. Region

- East Asia and Pacific
- South Asia
- Europe and Central Asia
- Latin America and Caribbean
- Middle East and North Africa
- Sub-Saharan Africa
- North America

10. Country, Region (within country) if applicable \*

11. Target Population Gender \*

*Multiple selection possible*

- Female
- Male

12. Target Population \*

- Individual
- Household
- Villages/Communities
- Subnational (district/state/county/etc.)
- National

13. Target Population Age \*

- Young Adults (18-35)
- Adults (36-65)
- Elderly (65+)
- Everyone
- Not specified

14. Target Population Living Environment \*

- Rural
- Urban
- Can't tell

15. Target population specific restrictions (e.g. pregnant women)

16. Sample Size of the Treatment Arm \*

Please enter a number greater than 0

17. Nature of comparison group \*

*Method of assignment to treatment or control*

- Randomized experiment
- Quasi-experiment, prospective assignment
- Quasi-experiment, ex-post assignment
- DiD
- Non-random
- Can't tell

18. Nature of comparison Group \*

*Unit of assignment (if applicable)*

- Individual
- Cluster
- Can't tell

19. Number of Clusters \*

The value must be a number

20. Outcome timing \*

*Time since the start of the intervention implementation*

- 1 to 3 years
- More than 3 years
- Can't tell

21. Number of relevant outcomes for this treatment arm \*

*Only behaviour outcomes concerning nutrition and diet, hygiene/WASH, or using health services.*

- 1
- 2
- 3
- 4
- 5
- 6
- 7
- 8
- 9
- 10

22. Notes

---

This content is neither created nor endorsed by Microsoft. The data you submit will be sent to the form owner.

 Microsoft Forms

# Data Extraction Public Health - Outcome

Alternative to EPPI to extract data from papers for the public health sector. This form is only for coding outcomes for which the forms for the general study and treatment arm have already been filled in. If you haven't filled in the form for the overall study data or the treatment arm for this outcome, please do that before starting with the outcome coding.

When some required data is unknown, please put in 7777.

\* Required

## General Information

1. ID of person extracting data \*

2. Year Of Publication \*

Number must be between 1980 ~ 2021

3. Item ID internal \*

*You can find this in EPPI (search by first author for speed)*

Number must be between 10000000 ~ 99999999

4. Which Treatment Arm does this Outcome belong to? \*

- treatment arm 1
- treatment arm 2
- treatment arm 3
- treatment arm 4
- treatment arm 5
- treatment arm 6
- treatment arm 7
- treatment arm 8
- treatment arm 9
- treatment arm 10

5. Which Outcome of the Treatment Arm you just specified are you coding? \*

- outcome 1
- outcome 2
- outcome 3
- outcome 4
- outcome 5
- outcome 6
- outcome 7
- outcome 8
- outcome 9
- outcome 10

6. Outcome Classification \*

- hygiene practices
- nutrition and dietary practices
- utilization of health services

7. Outcome \*

*note: HIV testing counts as medical check-up*

- anc
- expenditure
- food intake/preparation
- hand washing
- immunization/vaccination
- institutional delivery
- medical check-up
- medicine adherence
- open defecation
- pnc
- toilet construction
- toilet/latrine use
- water treatment

8. Description of the outcome \*

*This should be short, and only to differentiate between the other outcomes of the same intervention arm and study.*

9. Direction of the Effect \*

- Effect favours treatment
- Effect favours comparison
- Can't tell

10. What intervention did the comparison group receive

*If applicable*

- No treatment
- Treatment as usual
- Alternative intervention
- Other
- Can't tell

11. Were there any differences in the measurement of this outcome between the treatment group and the comparison group? \*

- Yes
- No
- Can't tell

12. Is the effect statistically significant? \*

- Yes
- No
- Can't tell



13. Significance level \*

- <0.10
- <0.05
- <0.01
- <0.001

14. Treatment Sample Size \*

The value must be a number

15. Control Sample Size \*

The value must be a number

16. Is the outcome from a regression model? \*

- Yes
- No

17. Nature of the Measures \*

- Continuous
- Dichotomous

Questions for Continuous Measures

18. Treatment Group Mean \*

The value must be a number

19. Comparison Group Mean \*

The value must be a number

20. Are the means reported above adjusted? \*

Yes

No

21. How is the variation expressed? \*

Standard Deviations of Treatment and Control

Standard Errors of Treatment and Control

t-test

Other

22. Treatment Group Standard Deviation \*

The value must be a number

23. Control Group Standard Deviation \*

The value must be a number

24. Treatment Group Standard Error \*

The value must be a number

25. Control Group Standard Error \*

The value must be a number

26. t-value from an independent t-test \*

The value must be a number

27. Other: value(s) \*

*If there are several, separate by semi-colon*

## Questions for Dichotomous Measures

28. How are the measures expressed? \*

*If you only have the first measure, then input 0 as the SE and add what is given instead (e.g., confidence interval) in the later question that allows for alternative measures*

- Number of participants that experienced change
- Proportion of participants that experienced change
- Odds Ratio + SE
- Logged Odds Ratio + SE
- Relative Risk (Risk Ratio) + SE
- Risk Difference + SE
- Chi-Squared
- 
- Other

29. Number of Participants in Treatment Group who experienced a change \*

The value must be a number

30. Number of Participants in Comparison Group who experienced a change \*

The value must be a number

31. Proportion of Participants in Treatment Group who experienced a change (if they name percentages, divide by 100) \*

Number must be between 0 ~ 1

32. Proportion of Participants in Comparison Group who experienced a change (if they name percentages, divide by 100) \*

Number must be between 0 ~ 1

33. Are the proportions above adjusted for pretest variables? \*

- Yes
- No

34. Odds Ratio \*

The value must be a number

35. The variation of the Odds ratio is expressed as... \*

- Standard Error
- 
- Other

36. Standard Error of Odds ratio \*

37. Value(s) of alternative measure of variation \*

*If there are multiple numbers, please separate by semi-colon. e.g. if the study reports a confidence interval, here you write " lower bound ; upper bound", where the lower and upper bounds are the reported numbers*

38. Is the Odds Ratio adjusted? \*

Yes

No

39. Logged Odds Ratio \*

The value must be a number

40. The variation of the Logged Odds Ratio is expressed as... \*

Standard Error

Other

41. Standard Error of Logged Odds Ratio \*

The value must be a number

42. Value(s) of alternative measure of variation \*

*If there are multiple numbers, please separate by semi-colon. e.g. if the study reports a confidence interval, here you write " lower bound ; upper bound", where the lower and upper bounds are the reported numbers*

43. Is the Logged Odds Ratio adjusted? \*

Yes

No

44. Relative Risk (Risk Ratio) \*

The value must be a number

45. The variation of the Relative Risk (Risk Ratio) is expressed as... \*

Standard Error

Other

46. Standard Error of Relative Risk (Risk Ratio) \*

The value must be a number

47. Value(s) of alternative measure of variation \*

*If there are multiple numbers, please separate by semi-colon. e.g. if the study reports a confidence interval, here you write " lower bound ; upper bound", where the lower and upper bounds are the reported numbers*

48. Is the Relative Risk (Risk Ratio) adjusted? \*

Yes

No

49. Risk Difference \*

The value must be a number

50. The variation of the Risk Difference is expressed as... \*

Standard Error

Other

51. Standard Error of Risk Difference \*

The value must be a number

52. Value(s) of alternative measure of variation \*

*If there are multiple numbers, please separate by semi-colon. e.g. if the study reports a confidence interval, here you write "lower bound ; upper bound", where the lower and upper bounds are the reported numbers*

53. Is the Risk Difference adjusted? \*

Yes

No

54. Other Measure: values for the alternative measurement specified \*

*Please note the results with clear explanation of which number is what, and separate different numbers by semi-colons. We need both the effect size and a measure of it's variation (SE, SD, CI...).*

55. Chi-squared \*

*with  $df=1$  (2 by 2 contingency table)*

The value must be a number

56. Econometric Model Used \*

- Binomial regression
- Cox regression
- Fixed Effects regression (FE)
- GEE
- GLM
- GLS
- IV OLS
- log-binomial regression
- Logistic regression
- OLS
- Random Effects regression (RE)

Other

57. Is the Intervention Variable a Dummy Variable \*

- Yes
- No

58. Is there clustering \*

*Are observations grouped into clusters, e.g. households clustered in villages. Search for whether authors mention clusters.*

- Yes
- No

59. Number of Clusters \*

The value must be a number

60. List the significant control variables used, separated by semi-colons \*

*You can also list on what pages you find them. We just need the names, not the values of their coefficients or anything like that.*

61. Number of Control Variables; Explanation. \*

*First type the number; add a semi-colon, then add the explanation. This is the number used in the regression of the outcome we're interested in, including those that aren't noted with coefficients. Also list the pages where you find them. How do you calculate the number?*

62. Sample Mean of Y (outcome variable)

The value must be a number

63. Standard Deviation of Y (outcome variable)

The value must be a number

64. Sample Mean of X

*X is the intervention variable (e.g. dummy for treatment group, participation in WASH intervention...)*

The value must be a number

65. Standard Deviation of X

*X is the intervention variable, e.g. electricity access*

66. Number of observations in the regression \*

The value must be a number

67. The coefficient of the intervention variable (beta) is \*

- Standardized
- Non-Standardized

68. The standardized beta-coefficient is \*

The value must be a number

69. The non-standardized beta-coefficient is \*

The value must be a number

70. The variation in the intervention coefficient beta is captured by \*

- A standard error
  - A t-statistic value
  - A p-value
  -
- Other

71. The standard error is for \*

- Standardized beta
- Non-Standardized beta

72. The value of the standard error of standardized beta is \*

The value must be a number

73. The value of the standard error of non-standardized beta is \*

The value must be a number

74. Are the standard errors clustered \*

Yes

No

75. At what level are the standard errors clustered? \*

76. Are the standard errors robust? \*

Yes

No

77. The t-statistic is for \*

Standardized beta

Non-Standardized beta

78. The value of the t-statistic of standardized beta is \*

The value must be a number

79. The value of the t-statistic of non-standardized beta is \*

The value must be a number

80. The p-value is for \*

Standardized beta

Non-Standardized beta

81. The value of the p-value of standardized beta is \*

The value must be a number

82. The value of the p-value of non-standardized beta is \*

The value must be a number



83. Other: \*

Please give the value(s) of the other measure of beta. If there are several, please separate by semi-colon

84. What are the Degrees of Freedom for the regression?

Only answer if the number is given

85. What (if any) is the type of  $R^2$  reported? \*

- Non-adjusted  $R^2$
- Adjusted  $R^2$
- Pseudo  $R^2$
- There is no  $R^2$  stated for the regression of interest

86. What is the value of the non-adjusted  $R^2$ ? \*

The value must be a number

87. What is the value of the adjusted  $R^2$ ? \*

The value must be a number

88. What is the value of the pseudo  $R^2$ ? \*

The value must be a number

89. What calculation method was used to get Pseudo  $R^2$ ?

90. Commentary on methods

91. Data Type \*

- Panel
- Cross-Section
- Time-Series