Duke | GLOBAL HEALTH INSTITUTE

Duke Department of Biostatistics & Bioinformatics
Duke University School of Medicine

Assessing the reporting of binary effect measures in CRTs: results from a crowd-sourced methods based review

Elizabeth L. Turner, Joanne McKenzie, Stephen Nash, Andrew Forbes, Karla Hemming on behalf of all participants of the crowd-sourced methods based reviews

With big thanks to co-authors from Duke University: Alyssa Platt, John Gallis and Kaitlin Tetreault and to Christina Easter from University of Birmingham

Background

EXAMPLE OF REPORTING OF BINARY OUTCOME IN A CRT

Background

- Cluster randomised trials (CRTs) used to evaluate complex & community-based interventions
- CRTs analysis methods more complex than for iRCTs
- Binary outcomes common in CRTs
- CRT analysis methods even more challenging for binary
- CONSORT statement on reporting of binary outcomes
  - Point 17b: “both relative and absolute”

Example – CRT with common binary outcome

RESEARCH ARTICLE

Improving rational use of ACTs through diagnosis-dependent subsidies: Evidence from a cluster-randomized controlled trial in western Kenya

Wendy Prudhomme O'Meara[1,2,3]*, Diana Menya[3], Jeremiah Laktabai[4,5], Alyssa Platt[2,6], Indrani Saran[2], Elisa Maffioli[2,7], Joseph Kipkoech[5], Manoj Mohanan[2,7,8], Elizabeth L. Turner[2,6]

Example – CRT with common binary outcome

PLOS | MEDICINE

RESEARCH ARTICLE

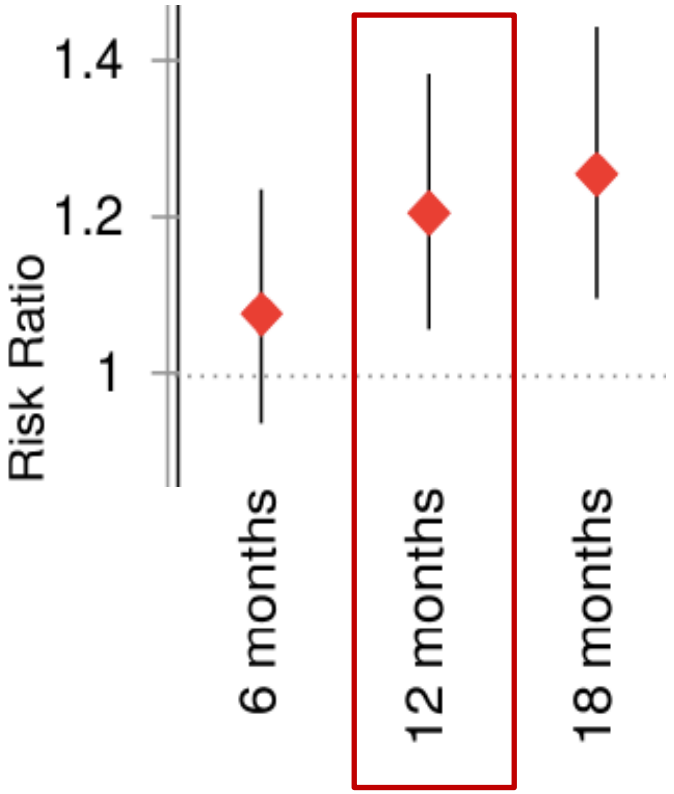
Improving rational use of ACTs through diagnosis-dependent subsidies: Evidence from a cluster-randomized controlled trial in western Kenya

Wendy Prudhomme O'Meara[1,2,3]*, Diana Menya[3], Jeremiah Laktabai[4,5], Alyssa Platt[2,6], Indrani Saran[2], Elisa Maffioli[2,7], Joseph Kipkoech[5], Manoj Mohanan[2,7,8], Elizabeth L. Turner[2,6]

- 32 communities
- FU: 6, 12 and 18 months
- ~2000 febrile participants/FU
- Primary outcome:
  - "Testing uptake" at 12 mths
- Common primary outcome:
  - 51% int. vs. 43% control

Example – CRT with common binary outcome

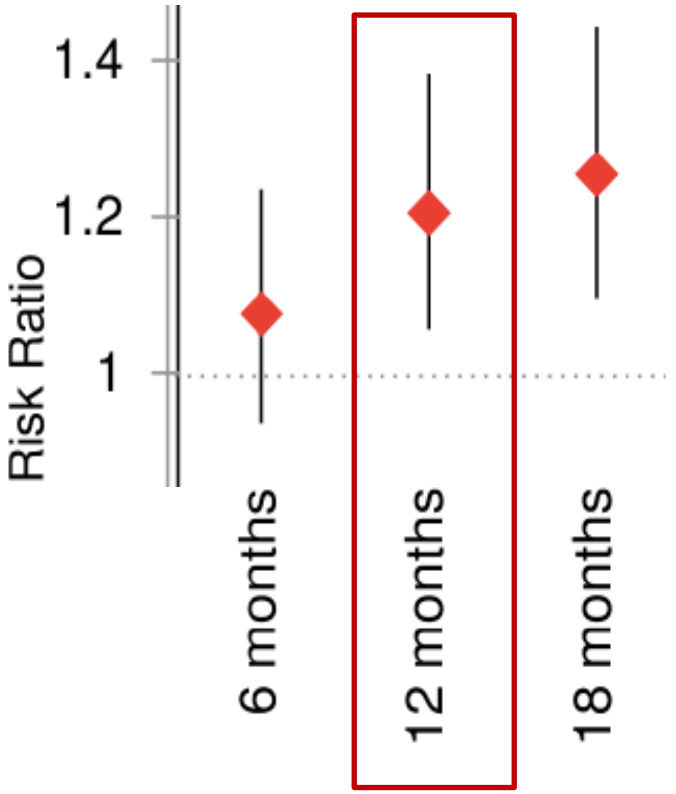
51% vs. 43%
at 12 months



Primary outcome: “testing uptake”

Example – CRT with common binary outcome

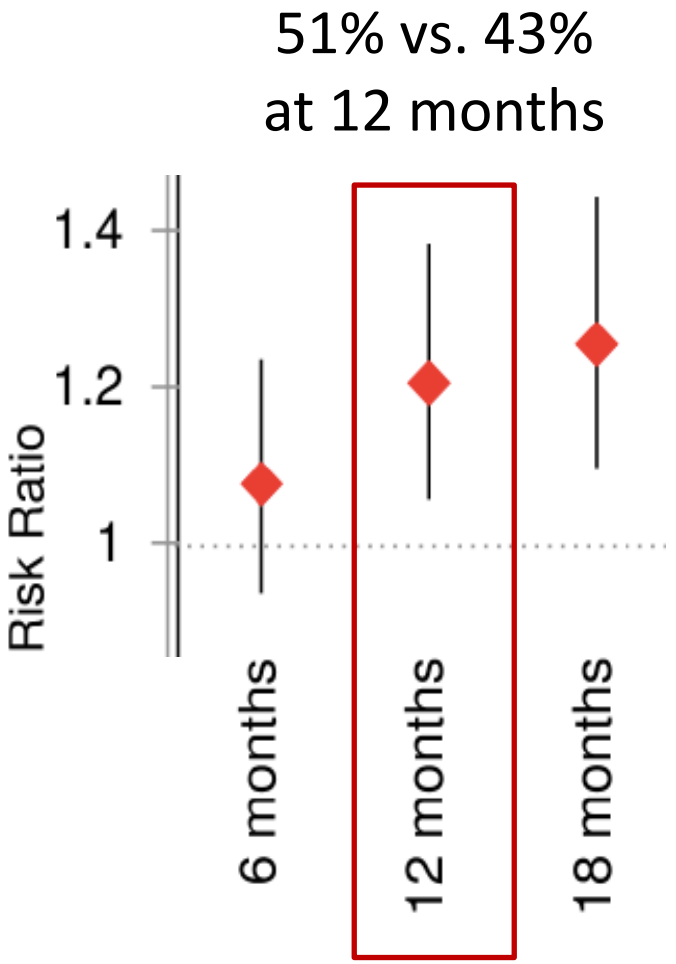
51% vs. 43%
at 12 months



Primary outcome: “testing uptake”

- Abstract
  - Both relative & absolute
  - Relative: risk ratio (RR)
  - Absolute: risk difference (RD)

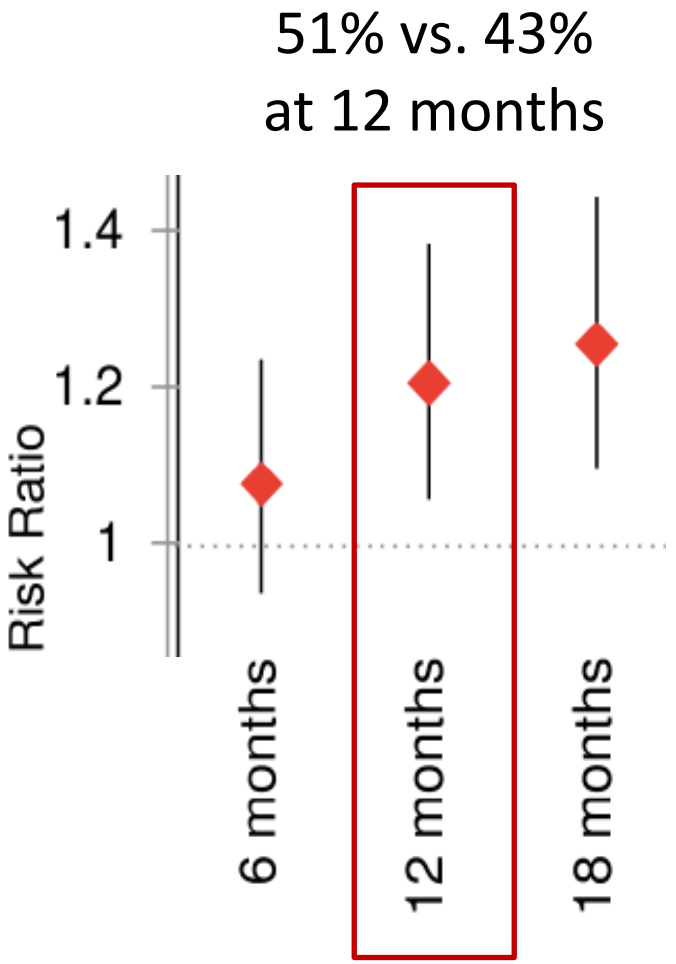
Example – CRT with common binary outcome



Primary outcome: “testing uptake”

- Abstract
  - Both relative & absolute
  - Relative: risk ratio (RR)
  - Absolute: risk difference (RD)
- Tables in main text
  - Both RR and RD

Example – CRT with common binary outcome

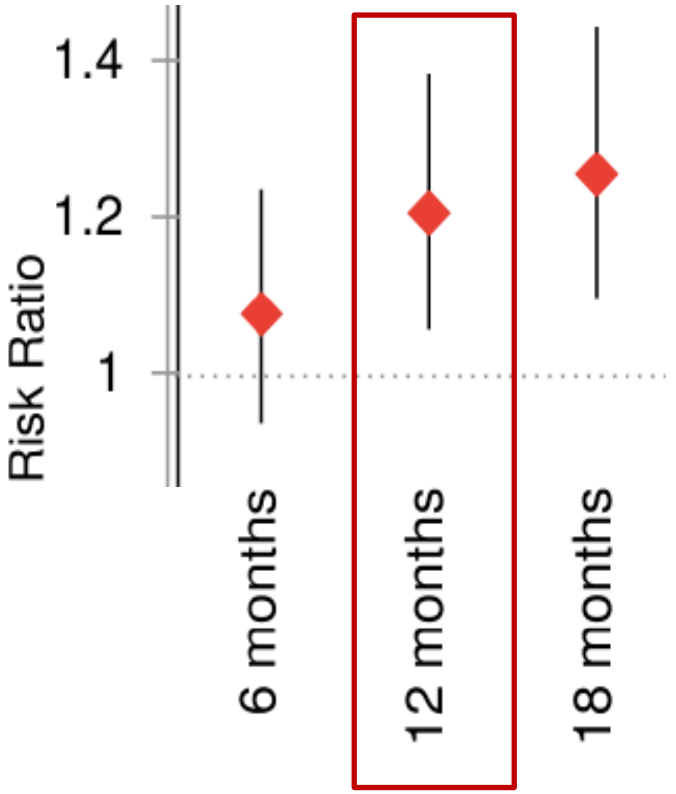


Primary outcome: “testing uptake”

- Abstract
  - Both relative & absolute
  - Relative: risk ratio (RR)
  - Absolute: risk difference (RD)
- Tables in main text
  - Both RR and RD
- Figures
  - RR only

Example – CRT with common binary outcome

51% vs. 43%
at 12 months

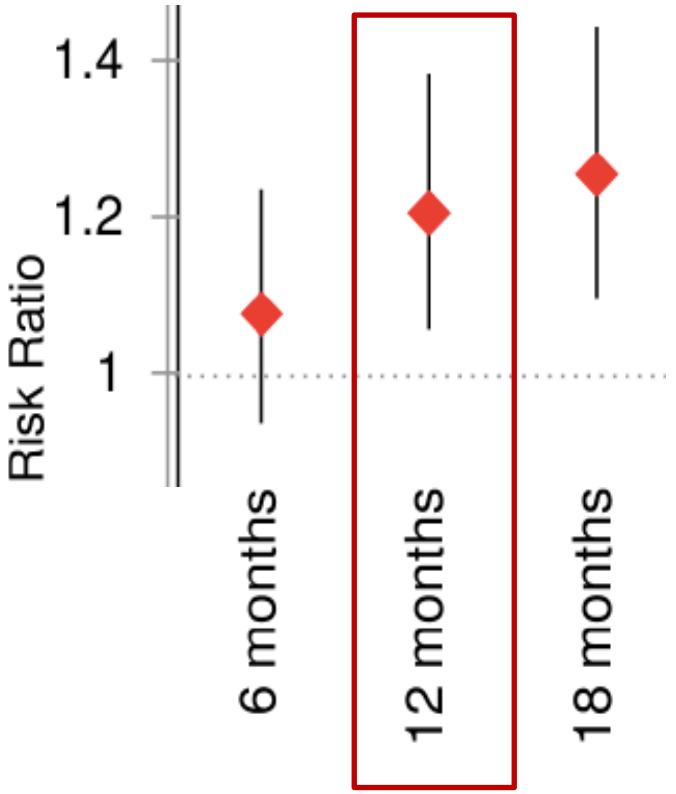


Primary outcome: “testing uptake”

- Common outcome
  - RD (95% CI): +9pp (+2,+15)
  - RR (95% CI): 1.20 (1.05,1.38)
  - OR > RR and approx. 1.49

Example – CRT with common binary outcome

51% vs. 43%
at 12 months



Primary outcome: “testing uptake”

- Common outcome
  - RD (95% CI): +9pp (+2,+15)
  - RR (95% CI): 1.20 (1.05,1.38)
  - OR > RR and approx. 1.49
- If interpreted OR as RR, would over-state magnitude of association of intervention and primary outcome

Goals of "crowd-sourced" methods review of binary outcomes in CRTs

1. Summarize effect measures for binary outcomes in CRTs
2. Compare to CONSORT recommendations
3. Summarize the statistical methods used
4. Identify opportunities to raise awareness of issues and to clarify methods to the community
5. Highlight the pros and cons of the "crowd-sourced" approach

Methods

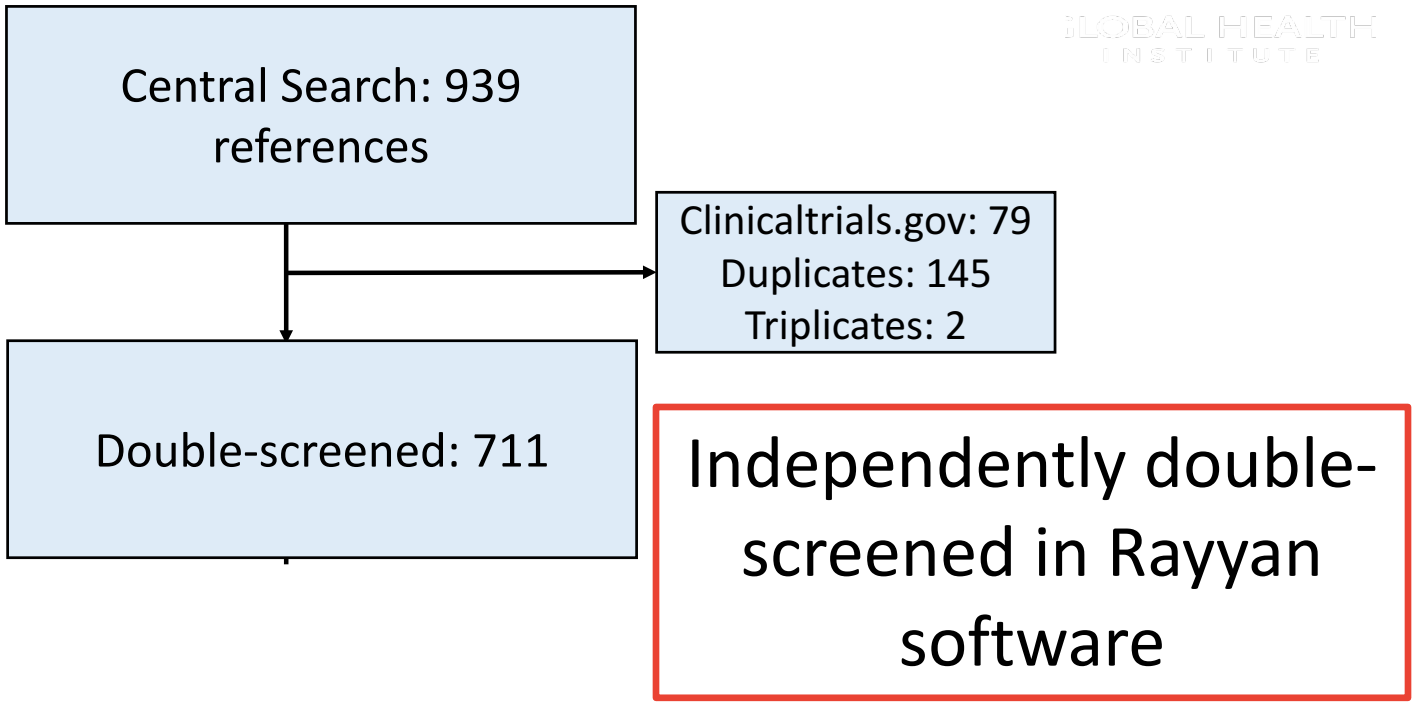
REVIEW OF REPORTING OF BINARY OUTCOMES IN CRTs

Methods – Inclusion criteria

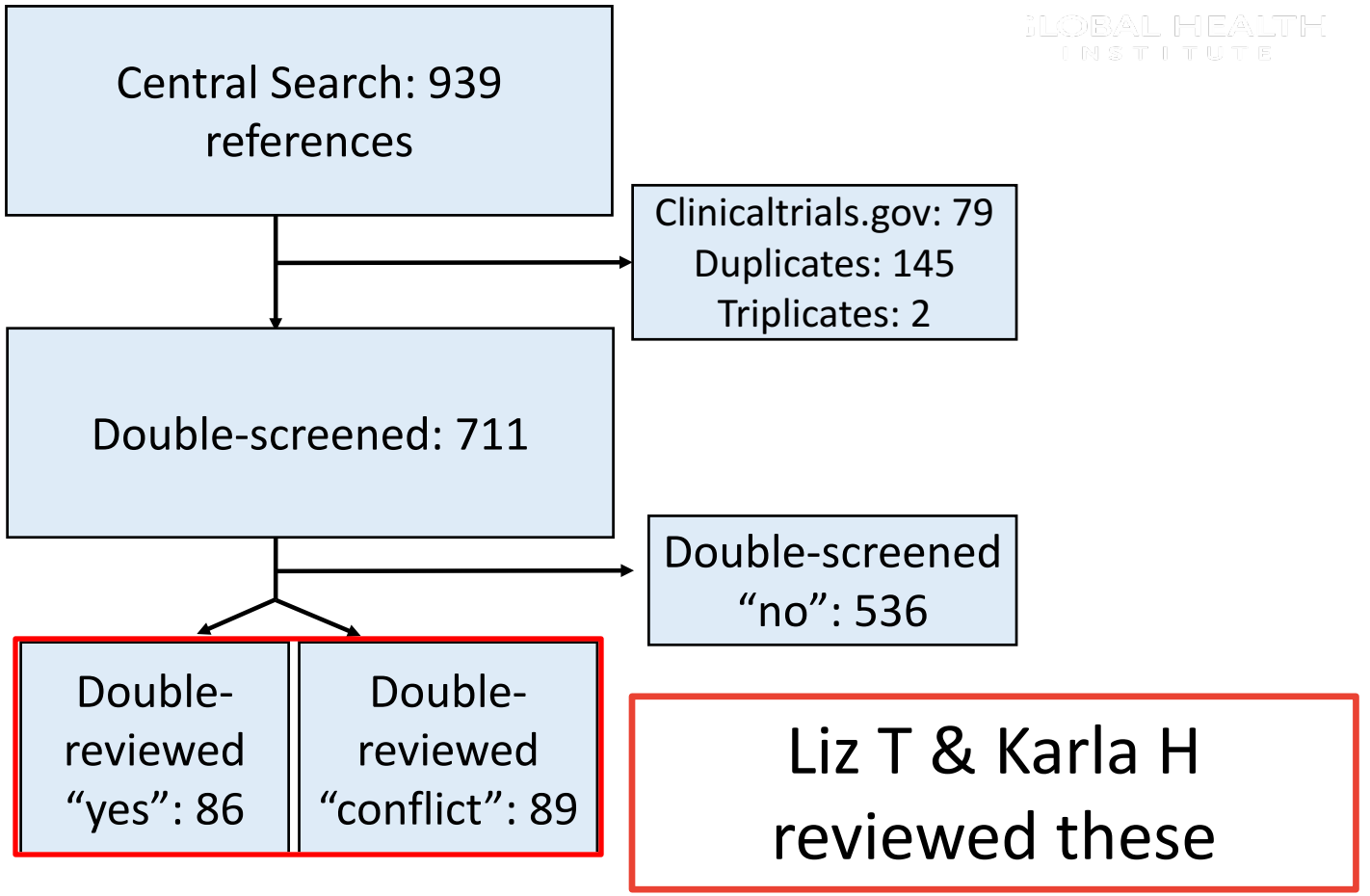
- Two-arm parallel CRT with ≥ 1 binary primary outcome
  - No stepped-wedge, crossover, factorial designs
- Main analysis of "definitive" CRT
  - No pilot/feasibility CRTs
  - No secondary or subgroup papers
- Peer-reviewed report in 2017 (either online or "in print")
  - No conference proceedings or only on trial registration website

Methods – Search strategy and papers identified

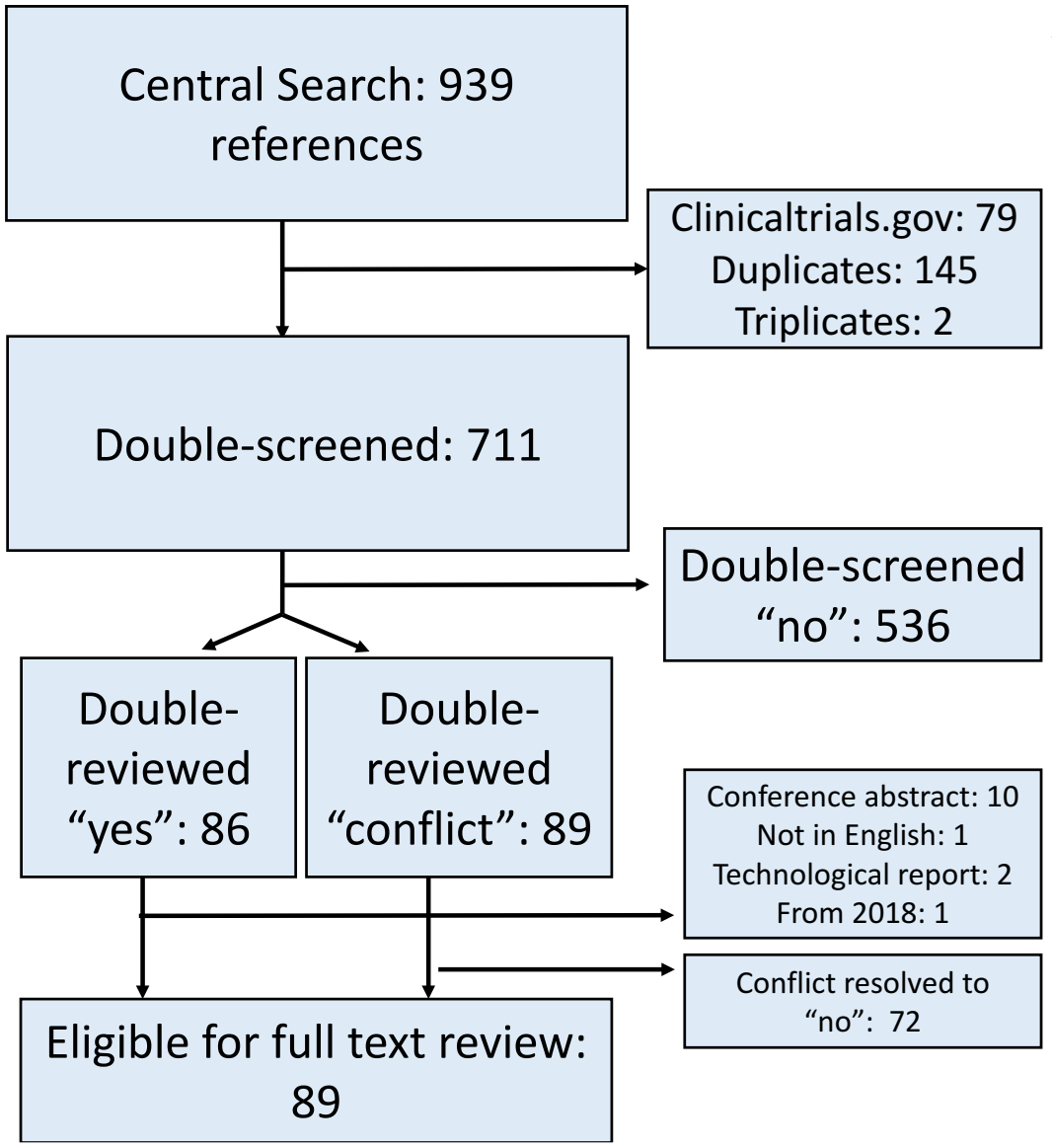
Methods – Search strategy and papers identified



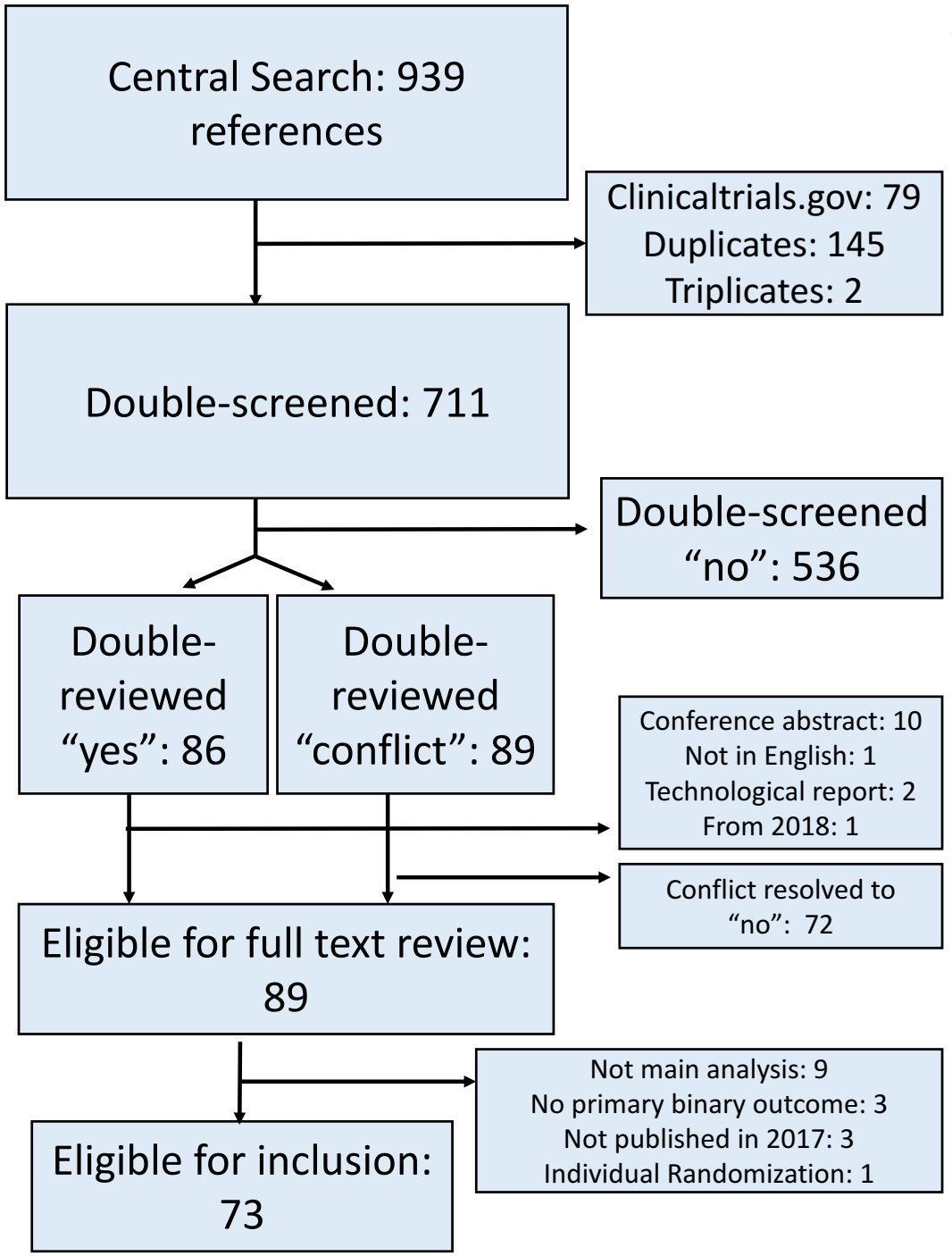
Methods – Search strategy and papers identified



Methods – Search strategy and papers identified

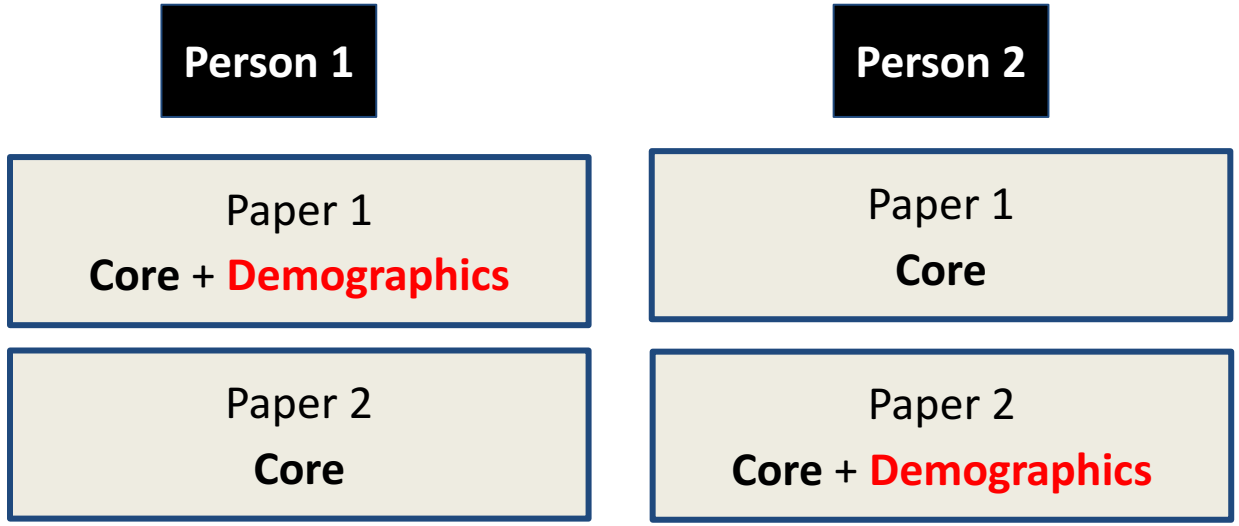


Methods – Search strategy and papers identified



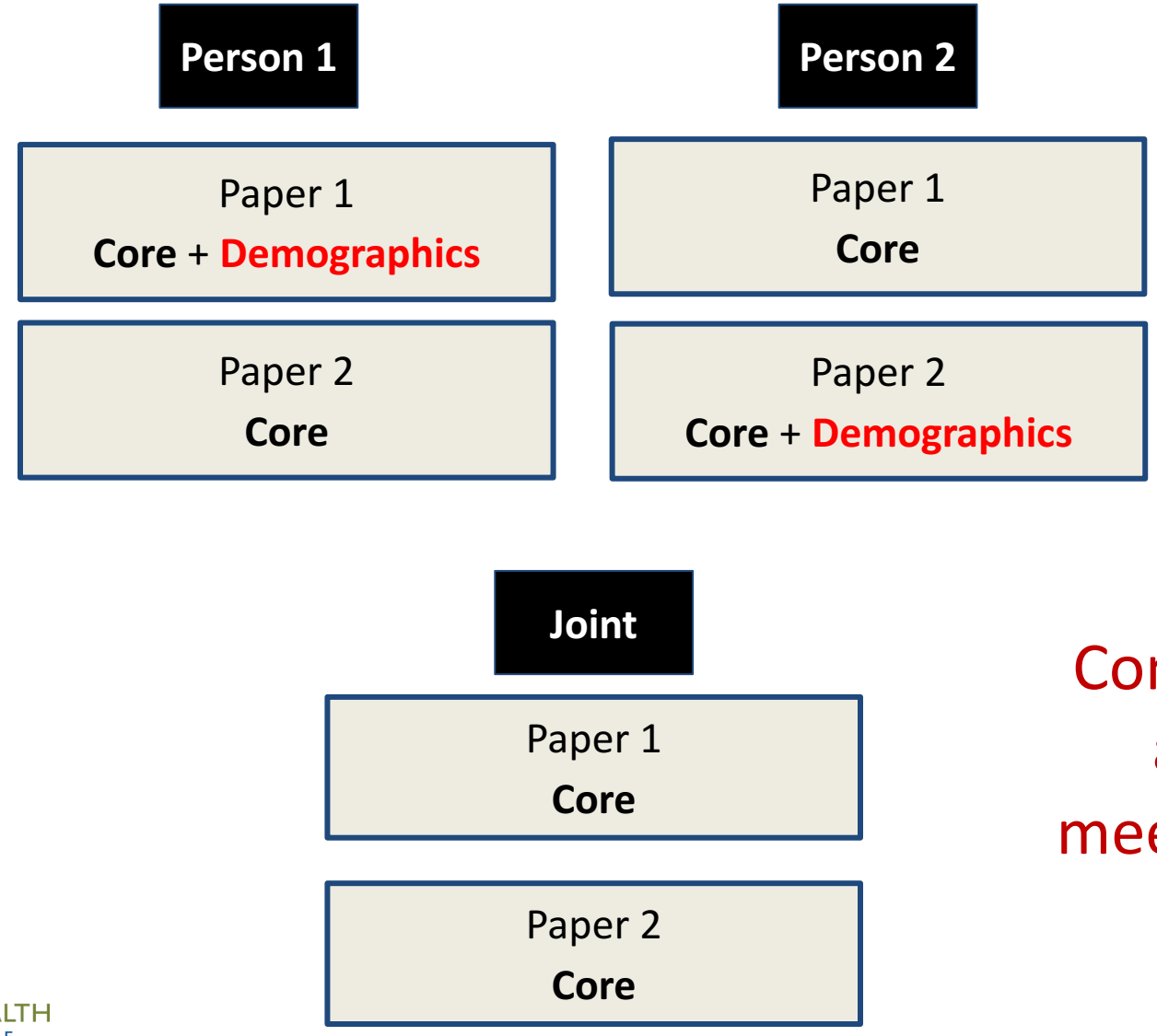
Methods – Data abstraction – Process for each article

Methods – Data abstraction – Process for each article



Completed independently online via Recap before in-person meeting

Methods – Data abstraction – Process for each article



Completed jointly at in-person meeting & entered in Redcap

Methods – Data reconciliation – "Crowd-sourcing"

- Three in-person meetings
  - Last year's QMUL conference (Nov '18)
  - Duke University Biostatistics Core (March '19)
  - University of Birmingham CTU (April '19)
- Initially planned only QMUL meeting
  - Added two extra to enable us to extract data from all N=73 articles
- Overall, 85 reviewers participated
  - Many of you here today
  - Thank you!

Methods – Challenges

- Unfunded project = difficulty protecting time
- Data cleaning
  - Skip patterns and comments
- Additional data extraction
  - Duke team extracted additional data from all 73 articles
    - Including whether CONSORT mentioned in CRT report

Results

REVIEW OF REPORTING OF BINARY OUTCOMES IN CRTs

Results - Reviewer Characteristics (N=85)

* Not mutually exclusive

Characteristic	N (%)
Highest career level	
Student (PhD/MSc)	35 (42.7%)
Researcher (Post-doc & MSc-level)	27 (32.9%)
University Prof.	20 (24.4%)
Main role	
Methodologist (Statistician)	68 (84%)
Trialist/Other	13 (16%)
Type of work setting*	
Healthcare	16 (18.8%)
University	71 (83.5%)
Country of work	
United Kingdom	51 (62.2%)
United States	19 (23.2%)
Other	12 (14.1%)
Previous CRT experience	
None	27 (32.9%)
1-2 CRT	28 (34.2%)
≥ 3 CRTs	27 (32.9%)

Results - Reviewer Characteristics (N=85)

* Not mutually exclusive

Characteristic	N (%)
Highest career level	
Student (PhD/MSc)	35 (42.7%)
Researcher (Post-doc & MSc-level)	27 (32.9%)
University Prof.	20 (24.4%)
Main role	
Methodologist (Statistician)	68 (84%)
Trialist/Other	13 (16%)
Type of work setting*	
Healthcare	16 (18.8%)
University	71 (83.5%)
Country of work	
United Kingdom	51 (62.2%)
United States	19 (23.2%)
Other	12 (14.1%)
Previous CRT experience	
None	27 (32.9%)
1-2 CRT	28 (34.2%)
≥ 3 CRTs	27 (32.9%)

Results - Reviewer Characteristics (N=85)

* Not mutually exclusive

Characteristic	N (%)
Highest career level	
Student (PhD/MSc)	35 (42.7%)
Researcher (Post-doc & MSc-level)	27 (32.9%)
University Prof.	20 (24.4%)
Main role	
Methodologist (Statistician)	68 (84%)
Trialist/Other	13 (16%)
Type of work setting*	
Healthcare	16 (18.8%)
University	71 (83.5%)
Country of work	
United Kingdom	51 (62.2%)
United States	19 (23.2%)
Other	12 (14.1%)
Previous CRT experience	
None	27 (32.9%)
1-2 CRT	28 (34.2%)
≥ 3 CRTs	27 (32.9%)

Results - Reviewer Characteristics (N=85)

* Not mutually exclusive

Characteristic	N (%)
Highest career level	
Student (PhD/MSc)	35 (42.7%)
Researcher (Post-doc & MSc-level)	27 (32.9%)
University Prof.	20 (24.4%)
Main role	
Methodologist (Statistician)	68 (84%)
Trialist/Other	13 (16%)
Type of work setting*	
Healthcare	16 (18.8%)
University	71 (83.5%)
Country of work	
United Kingdom	51 (62.2%)
United States	19 (23.2%)
Other	12 (14.1%)
Previous CRT experience	
None	27 (32.9%)
1-2 CRT	28 (34.2%)
≥ 3 CRTs	27 (32.9%)

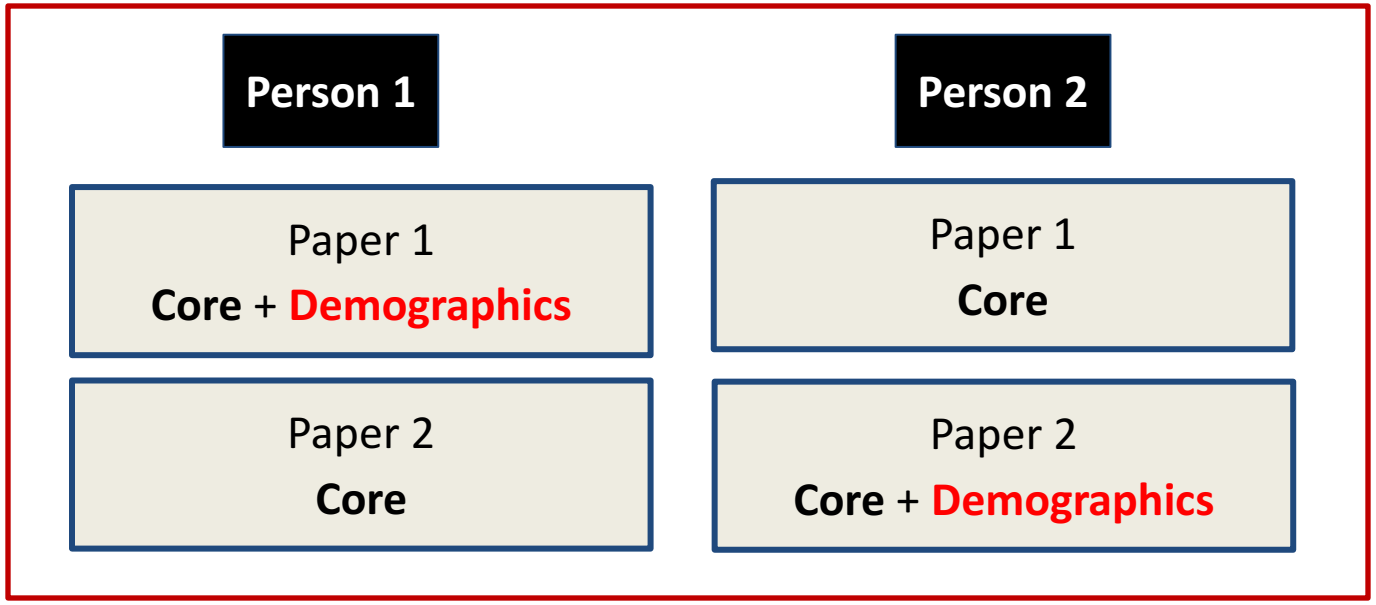
Results - Reviewer Characteristics (N=85)

* Not mutually exclusive

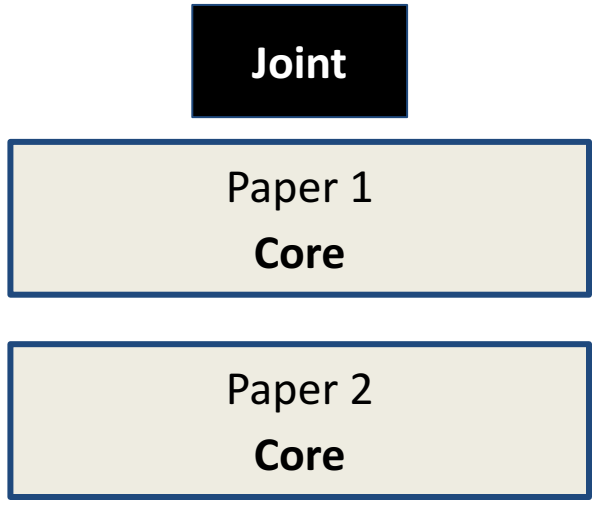
Characteristic	N (%)
Highest career level	
Student (PhD/MSc)	35 (42.7%)
Researcher (Post-doc & MSc-level)	27 (32.9%)
University Prof.	20 (24.4%)
Main role	
Methodologist (Statistician)	68 (84%)
Trialist/Other	13 (16%)
Type of work setting*	
Healthcare	16 (18.8%)
University	71 (83.5%)
Country of work	
United Kingdom	51 (62.2%)
United States	19 (23.2%)
Other	12 (14.1%)
Previous CRT experience	
None	27 (32.9%)
1-2 CRT	28 (34.2%)
≥ 3 CRTs	27 (32.9%)

Results – Agreement for common items across N=73 articles

Results – Agreement for common items across N=73 articles



85% agreement of 28 items in common across all articles and all participants



Results -
CRT Basic Characteristics (N=73)

* Not mutually exclusive

Results - CRT Basic Characteristics (N=73)

* Not mutually exclusive

Characteristic	N(%)
Three most common disease or domain under study*	
Infectious diseases	19 (26.0%)
Women's health	16 (21.9%)
Mental health and behavioural conditions	11 (15.1%)
Four most common geographic regions*	
Europe	22 (30.1%)
Africa	19 (26.0%)
Asia	14 (19.2%)
North America	12 (16.4%)
Most common type of experimental intervention*	
Direct participant therapeutic intervention	46 (64.8%)
Targeted at health care professionals	32 (44.4%)
Participant health promotion or educ. intervention	12 (16.9%)
Most common type of control Intervention	
Placebo, no active intervention	54 (74.0%)
Minimal application of experimental intervention	8 (11.0%)
Other	10 (13.7%)

Results - CRT Basic Characteristics (N=73)

* Not mutually exclusive

Characteristic	N(%)
Three most common disease or domain under study*	
Infectious diseases	19 (26.0%)
Women's health	16 (21.9%)
Mental health and behavioural conditions	11 (15.1%)
Four most common geographic regions*	
Europe	22 (30.1%)
Africa	19 (26.0%)
Asia	14 (19.2%)
North America	12 (16.4%)
Most common type of experimental intervention*	
Direct participant therapeutic intervention	46 (64.8%)
Targeted at health care professionals	32 (44.4%)
Participant health promotion or educ. intervention	12 (16.9%)
Most common type of control Intervention	
Placebo, no active intervention	54 (74.0%)
Minimal application of experimental intervention	8 (11.0%)
Other	10 (13.7%)

Results - CRT Basic Characteristics (N=73)

* Not mutually exclusive

Characteristic	N(%)
Three most common disease or domain under study*	
Infectious diseases	19 (26.0%)
Women's health	16 (21.9%)
Mental health and behavioural conditions	11 (15.1%)
Four most common geographic regions*	
Europe	22 (30.1%)
Africa	19 (26.0%)
Asia	14 (19.2%)
North America	12 (16.4%)
Most common type of experimental intervention*	
Direct participant therapeutic intervention	46 (64.8%)
Targeted at health care professionals	32 (44.4%)
Participant health promotion or educ. intervention	12 (16.9%)
Most common type of control Intervention	
Placebo, no active intervention	54 (74.0%)
Minimal application of experimental intervention	8 (11.0%)
Other	10 (13.7%)

Results - CRT Basic Characteristics (N=73)

* Not mutually exclusive

Characteristic	N(%)
Three most common disease or domain under study*	
Infectious diseases	19 (26.0%)
Women's health	16 (21.9%)
Mental health and behavioural conditions	11 (15.1%)
Four most common geographic regions*	
Europe	22 (30.1%)
Africa	19 (26.0%)
Asia	14 (19.2%)
North America	12 (16.4%)
Most common type of experimental intervention*	
Direct participant therapeutic intervention	46 (64.8%)
Targeted at health care professionals	32 (44.4%)
Participant health promotion or educ. intervention	12 (16.9%)
Most common type of control Intervention	
Placebo, no active intervention	54 (74.0%)
Minimal application of experimental intervention	8 (11.0%)
Other	10 (13.7%)

Results - CRT Design Characteristics (N=73)

* Not mutually exclusive

Results - CRT Design Characteristics (N=73)

* Not mutually exclusive

Characteristic	N(%)
Unit of Randomization	
Health facility	30 (41.1%)
Geographic areas (e.g. village or county)	14 (19.2%)
Health care provider	11 (15.1%)
School, School district	10 (13.7%)
Total Number of Clusters Randomized	
Median (25th, 75th percentile)	29 (20,44)
<6	3 (4.1%)
6-10	5 (6.8%)
11-20	17 (23.3%)
21-40	27 (37.0%)
>40	21 (28.8%)
Size of Average Clusters, Med(25th, 75th per.)	48 (20,220)
Min, Max	4, 9065
Study Design, n(%)	
Cohort	49 (67.1%)
Cross-sectional	23 (31.5%)
A mix of cohort, cross-sectional	1 (1.4%)

Results - CRT Design Characteristics (N=73)

* Not mutually exclusive

Characteristic	N(%)
Unit of Randomization	
Health facility	30 (41.1%)
Geographic areas (e.g. village or county)	14 (19.2%)
Health care provider	11 (15.1%)
School, School district	10 (13.7%)
Total Number of Clusters Randomized	
Median (25th, 75th percentile)	29 (20,44)
<6	3 (4.1%)
6-10	5 (6.8%)
11-20	17 (23.3%)
21-40	27 (37.0%)
>40	21 (28.8%)
Size of Average Clusters, Med(25th, 75th per.)	48 (20,220)
Min, Max	4, 9065
Study Design, n(%)	
Cohort	49 (67.1%)
Cross-sectional	23 (31.5%)
A mix of cohort, cross-sectional	1 (1.4%)

Results - CRT Design Characteristics (N=73)

* Not mutually exclusive

Characteristic	N(%)
Unit of Randomization	
Health facility	30 (41.1%)
Geographic areas (e.g. village or county)	14 (19.2%)
Health care provider	11 (15.1%)
School, School district	10 (13.7%)
Total Number of Clusters Randomized	
Median (25th, 75th percentile)	29 (20,44)
<6	3 (4.1%)
6-10	5 (6.8%)
11-20	17 (23.3%)
21-40	27 (37.0%)
>40	21 (28.8%)
Size of Average Clusters, Med(25th, 75th per.)	48 (20,220)
Min, Max	4, 9065
Study Design, n(%)	
Cohort	49 (67.1%)
Cross-sectional	23 (31.5%)
A mix of cohort, cross-sectional	1 (1.4%)

Results - CRT Design Characteristics (N=73)

* Not mutually exclusive

Characteristic	N(%)
Unit of Randomization	
Health facility	30 (41.1%)
Geographic areas (e.g. village or county)	14 (19.2%)
Health care provider	11 (15.1%)
School, School district	10 (13.7%)
Total Number of Clusters Randomized	
Median (25th, 75th percentile)	29 (20,44)
<6	3 (4.1%)
6-10	5 (6.8%)
11-20	17 (23.3%)
21-40	27 (37.0%)
>40	21 (28.8%)
Size of Average Clusters, Med(25th, 75th per.)	48 (20,220)
Min, Max	4, 9065
Study Design, n(%)	
Cohort	49 (67.1%)
Cross-sectional	23 (31.5%)
A mix of cohort, cross-sectional	1 (1.4%)

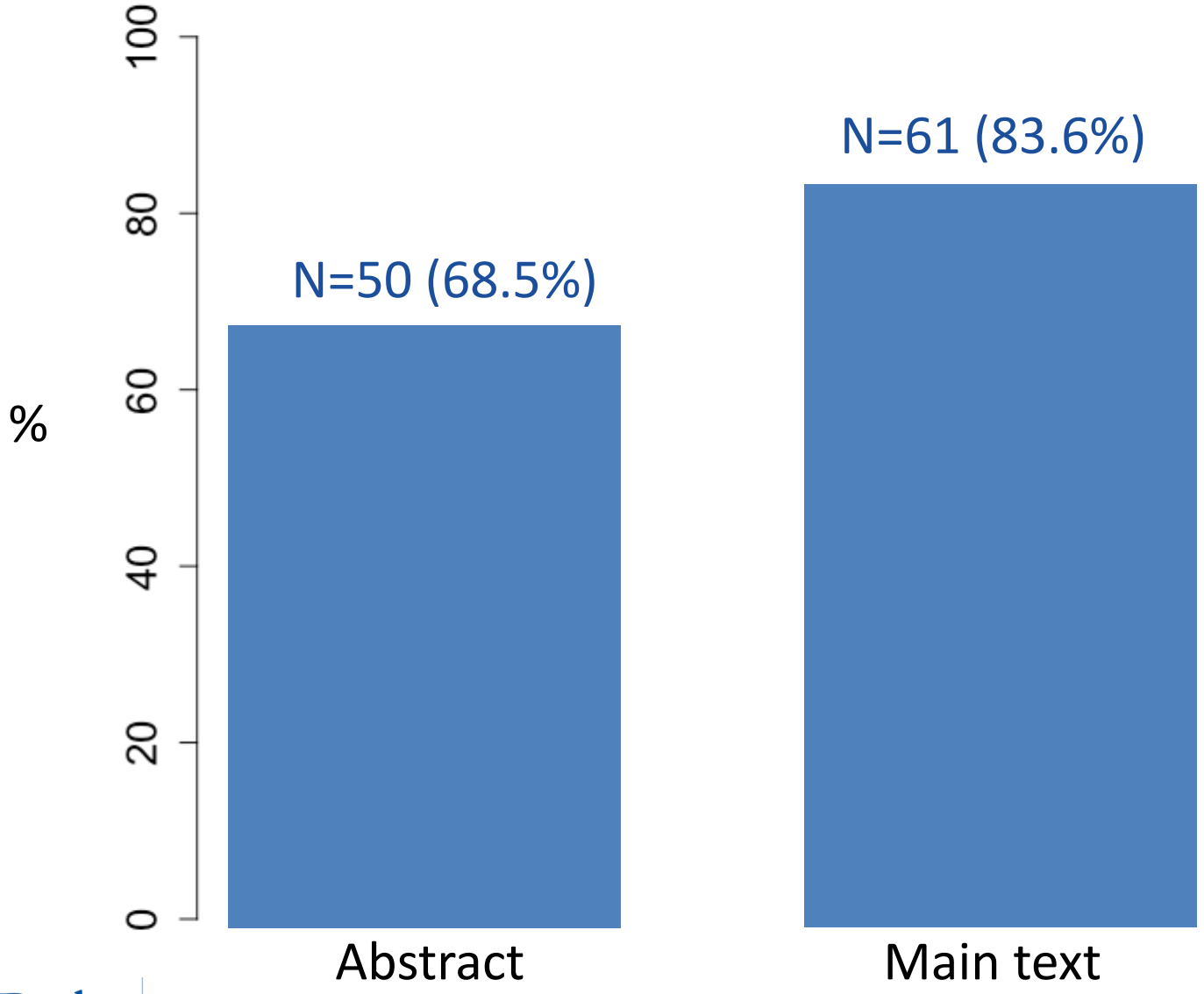
Results - CRT Design Characteristics (N=73)

* Not mutually exclusive

Characteristic	N(%)
Unit of Randomization	
Health facility	30 (41.1%)
Geographic areas (e.g. village or county)	14 (19.2%)
Health care provider	11 (15.1%)
School, School district	10 (13.7%)
Total Number of Clusters Randomized	
Median (25th, 75th percentile)	29 (20,44)
<6	3 (4.1%)
6-10	5 (6.8%)
11-20	17 (23.3%)
21-40	27 (37.0%)
>40	21 (28.8%)
Size of Average Clusters, Med(25th, 75th per.)	48 (20,220)
Min, Max	4, 9065
Study Design, n(%)	
Cohort	49 (67.1%)
Cross-sectional	23 (31.5%)
A mix of cohort, cross-sectional	1 (1.4%)

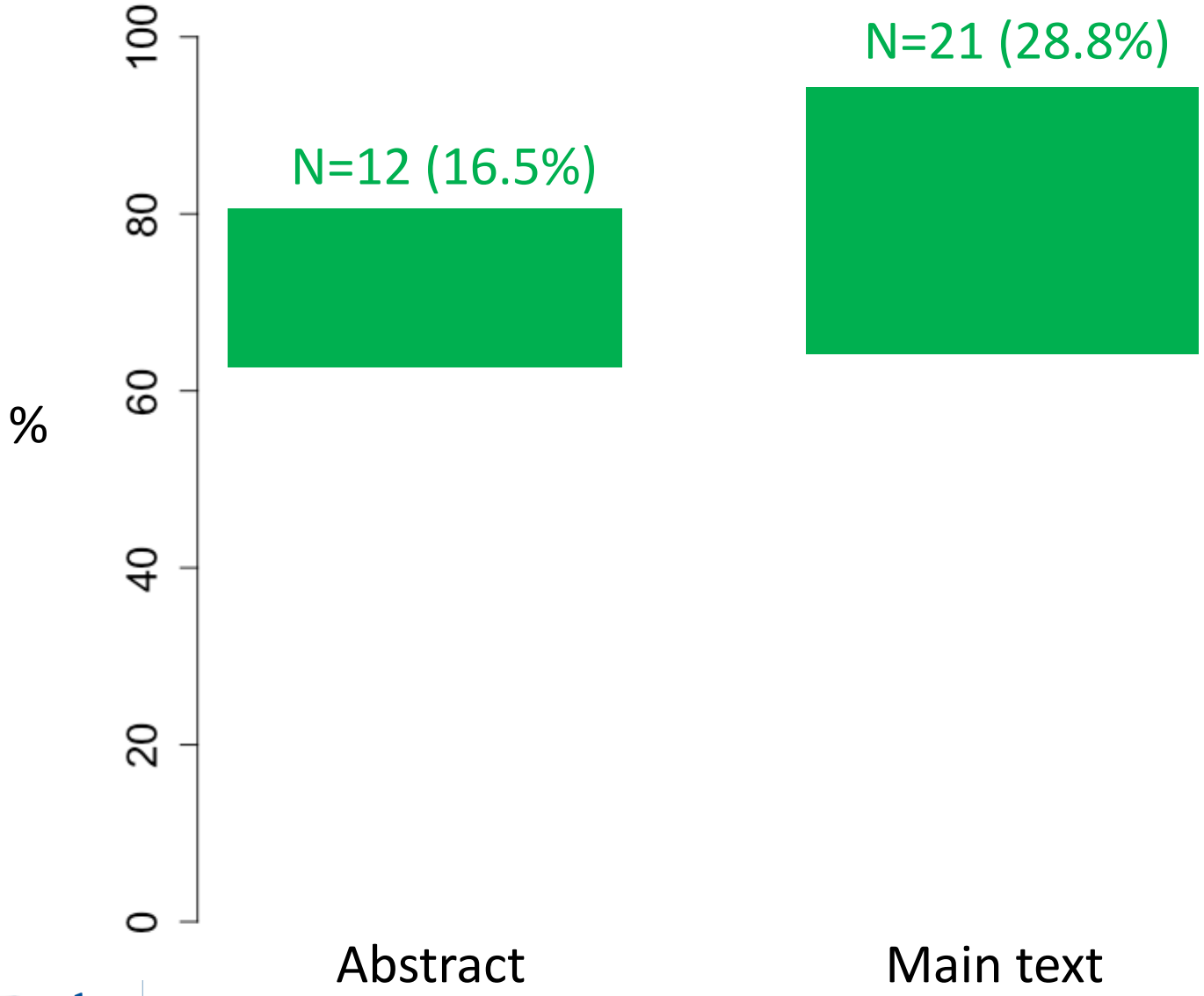
Results – Reporting of binary outcomes CRTs (N=73)

Results – Reporting of binary outcomes CRTs (N=73)



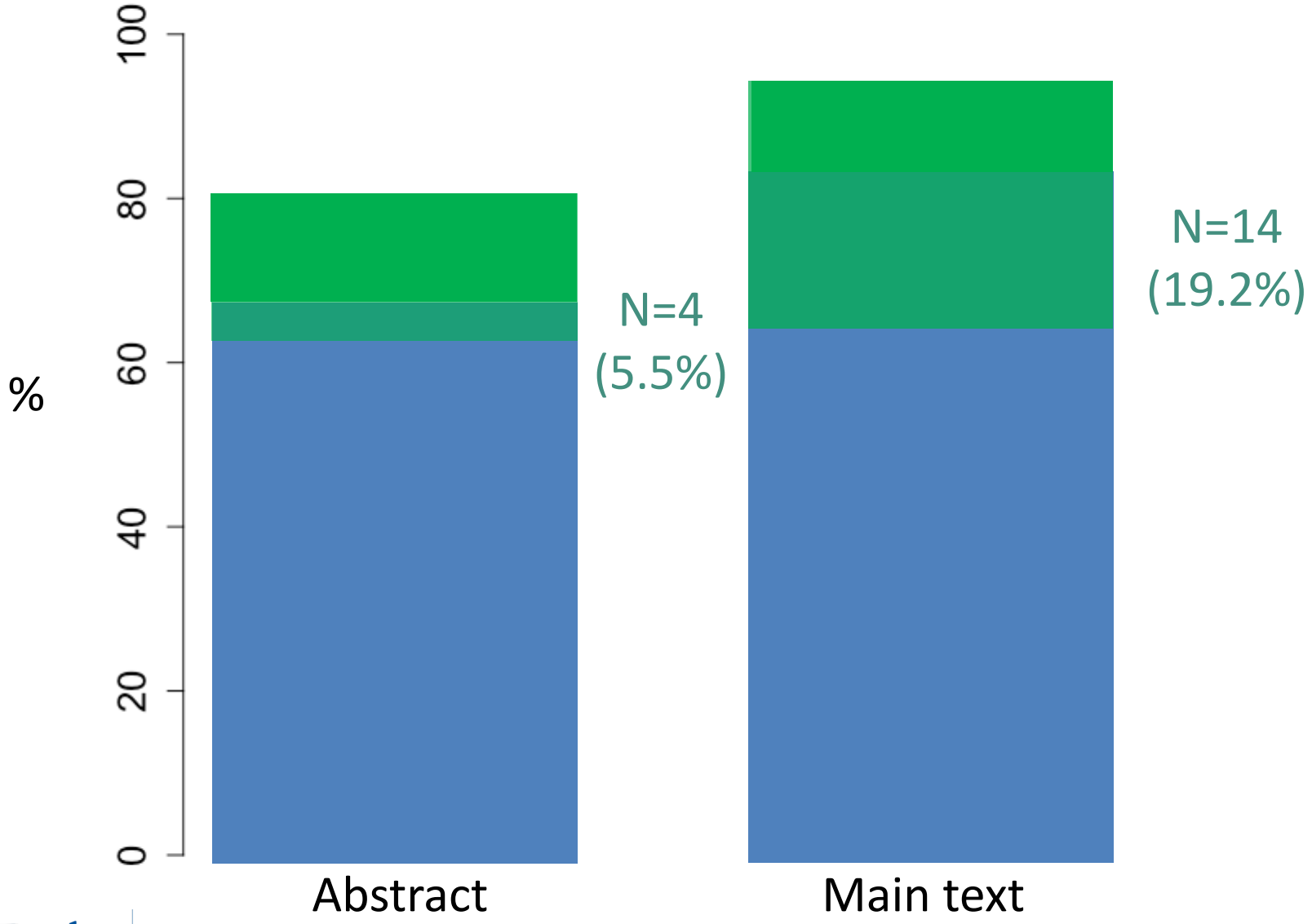
Relative effects

Results – Reporting of binary outcomes CRTs (N=73)



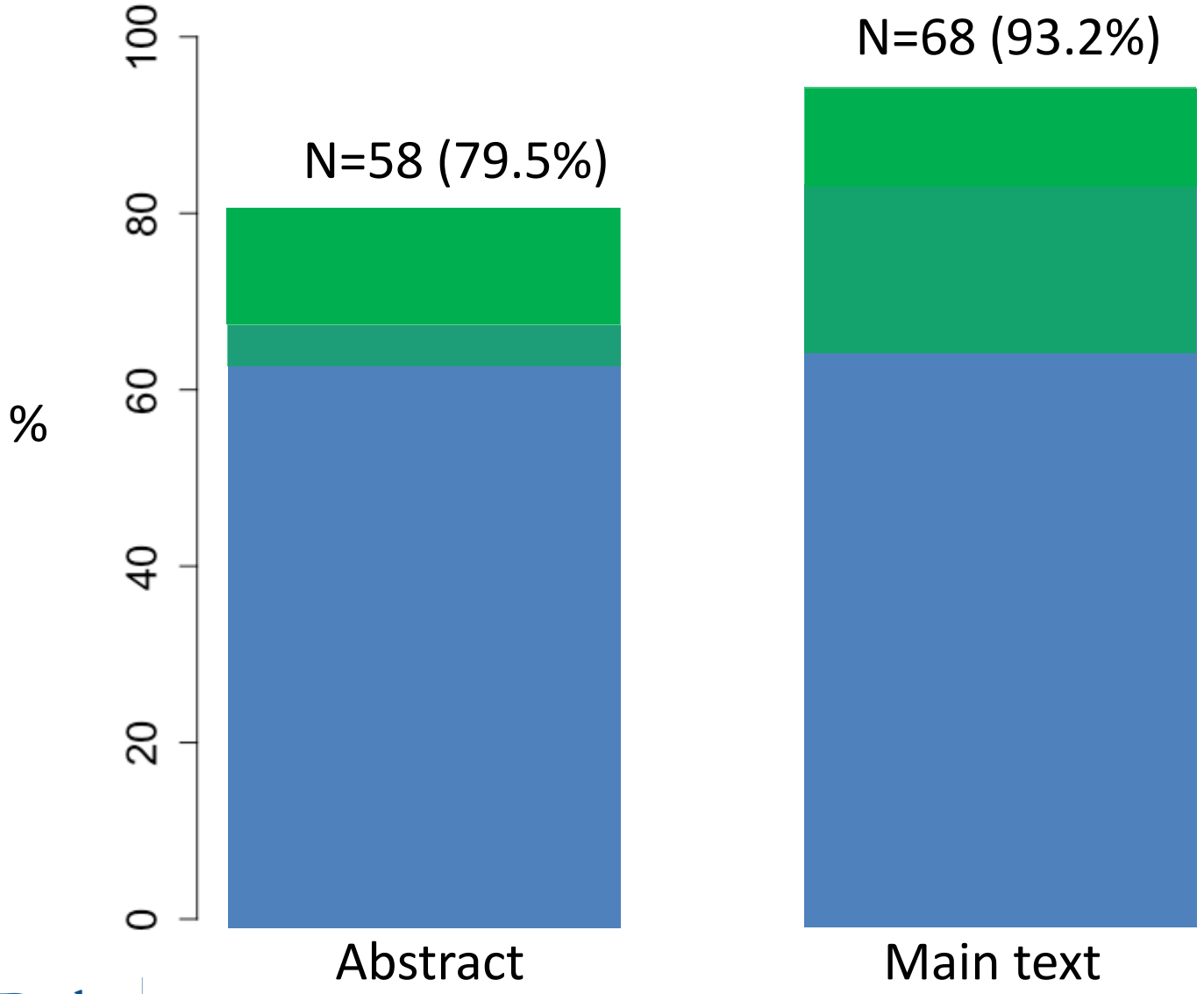
Absolute effects

Results – Reporting of binary outcomes CRTs (N=73)



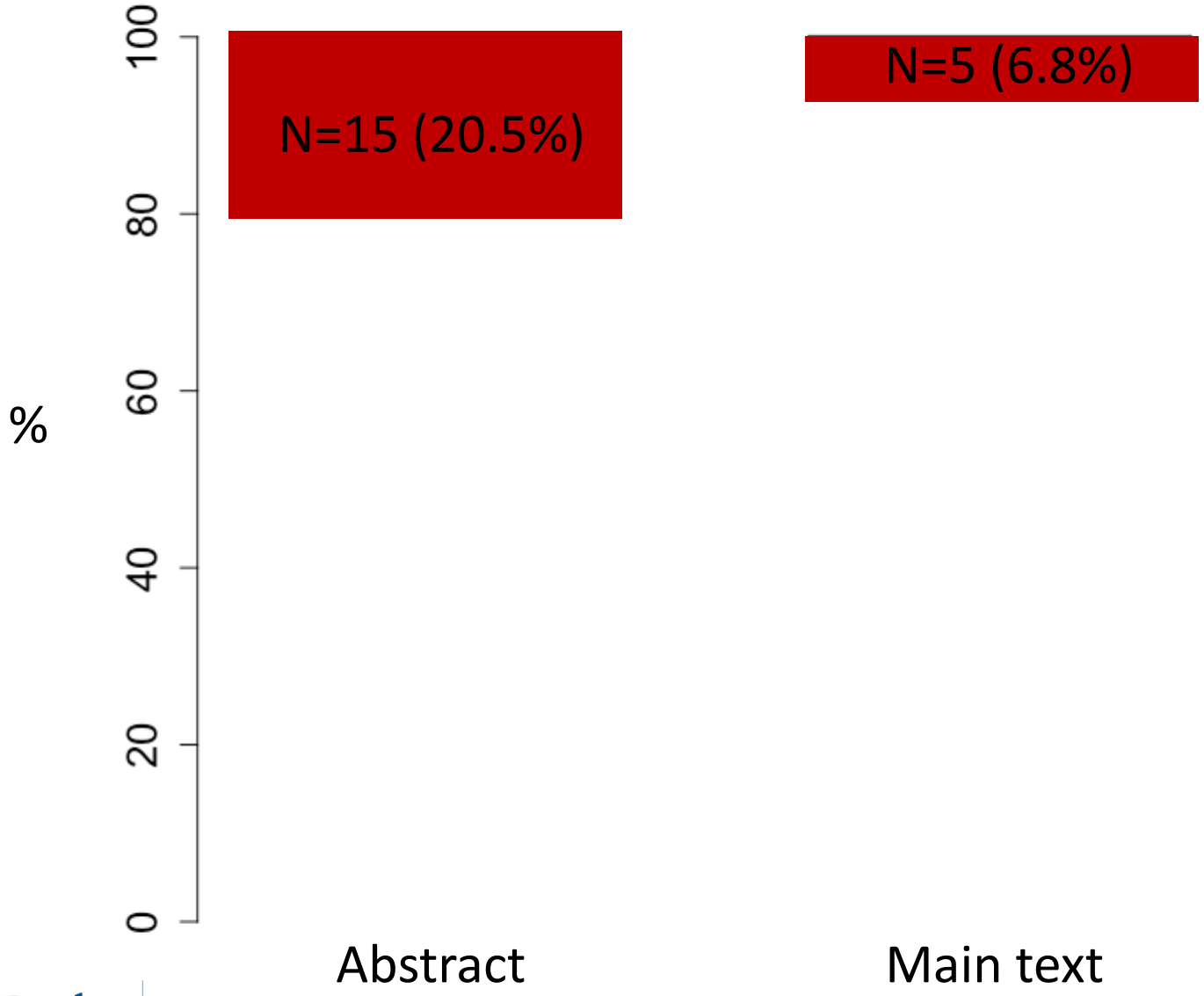
Both relative & absolute

Results – Reporting of binary outcomes CRTs (N=73)



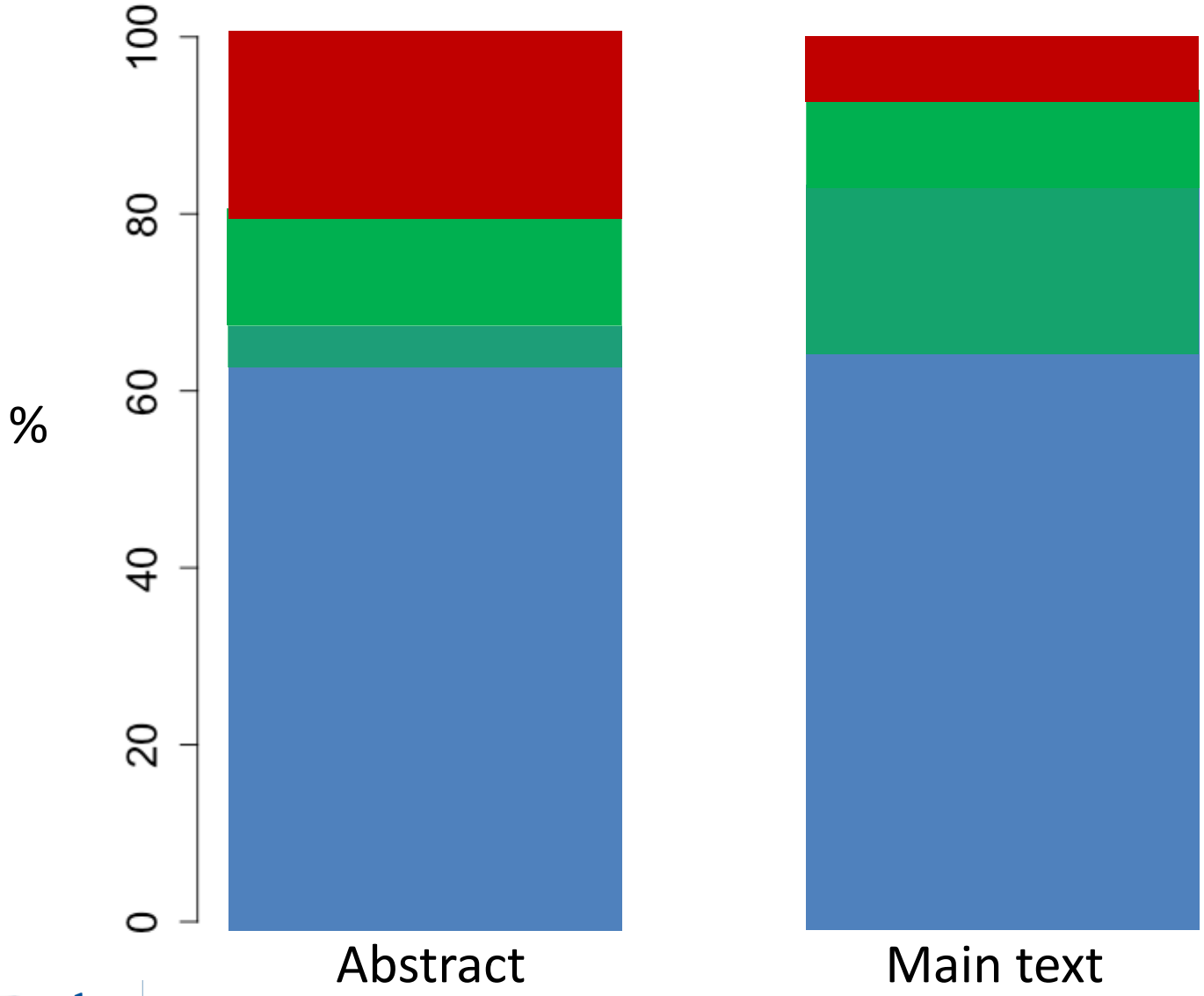
Some effect – either relative or absolute

Results – Reporting of binary outcomes CRTs (N=73)



No effect measure – neither relative nor absolute

Results – Reporting of binary outcomes CRTs (N=73)



No effect measure

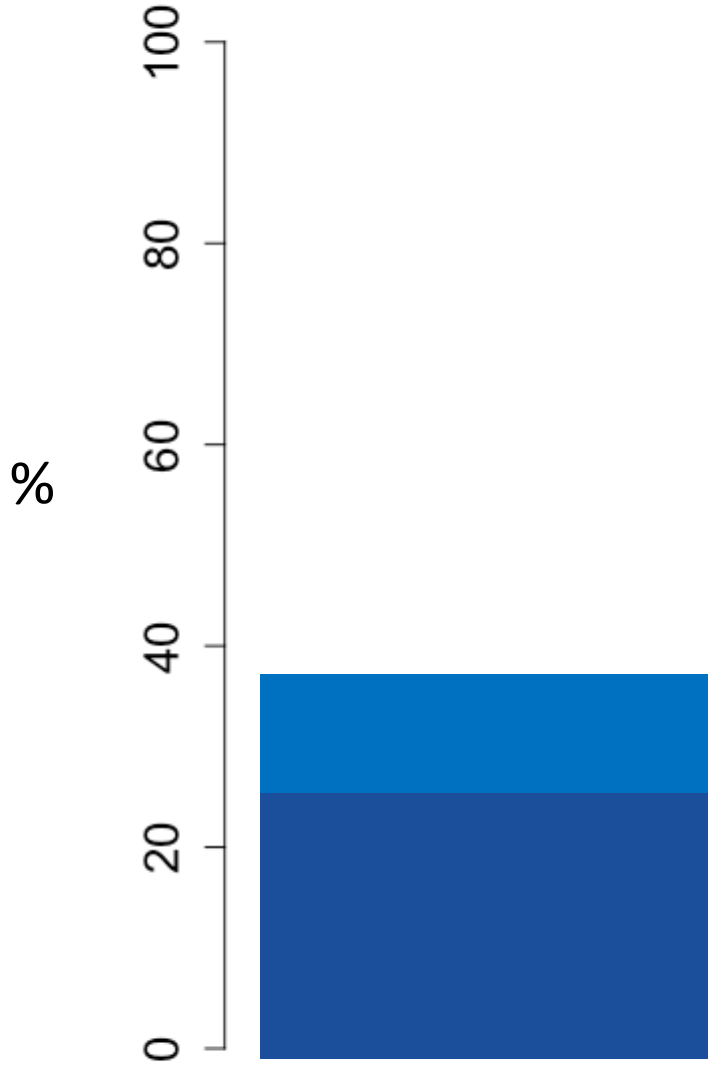
Absolute effects

Both relative & absolute

Relative effects

Results – Reference to CONSORT (N=73)

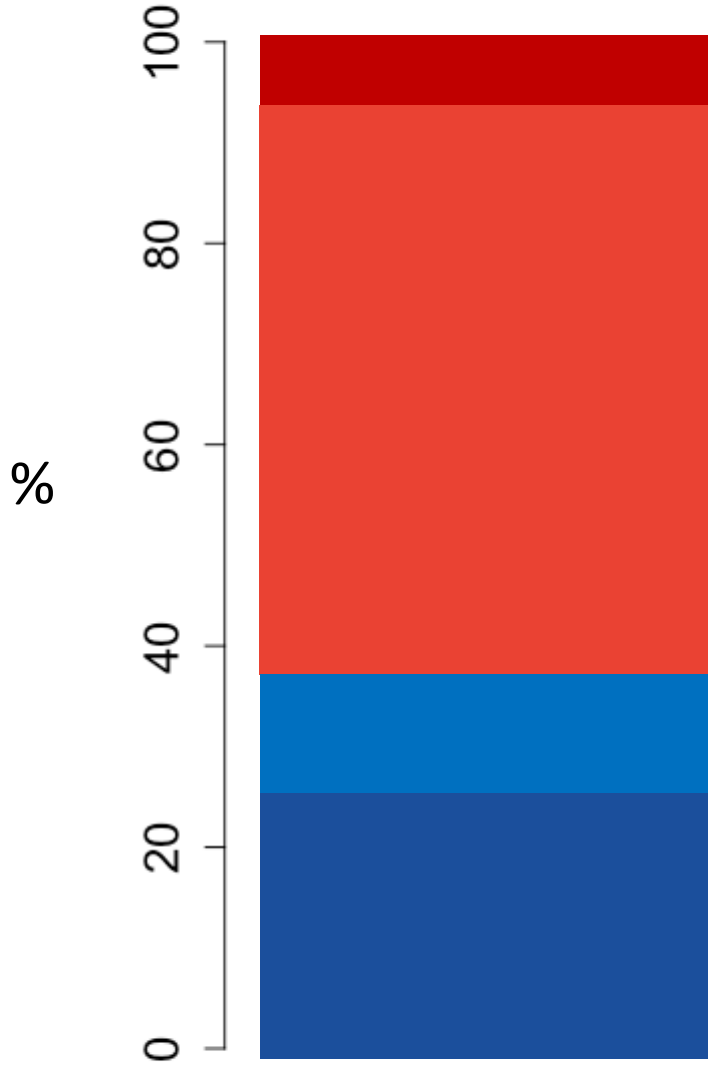
Results – Reference to CONSORT (N=73)



General CONSORT: N=9 (12.3%)

CONSORT extension for CRTs: N=18 (24.7%)

Results – Reference to CONSORT (N=73)



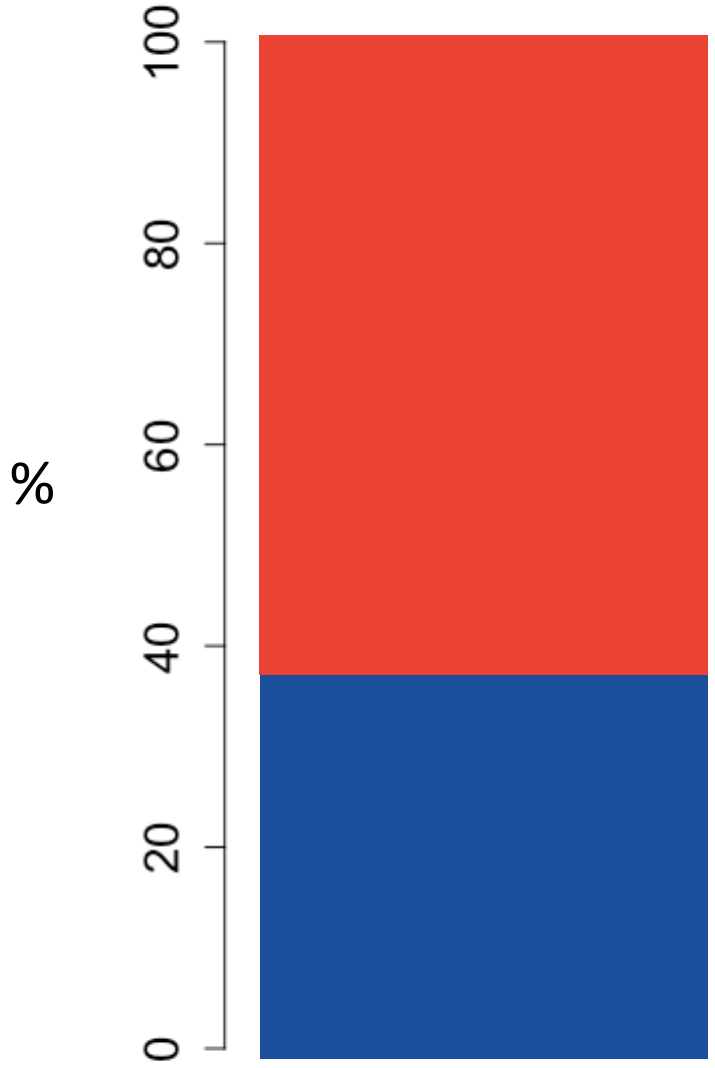
No mention of CONSORT: N=3 (4.1%)

CONSORT flow-chart only: N=43 (58.9%)

General CONSORT: N=9 (12.3%)

CONSORT extension for CRTs: N=18 (24.7%)

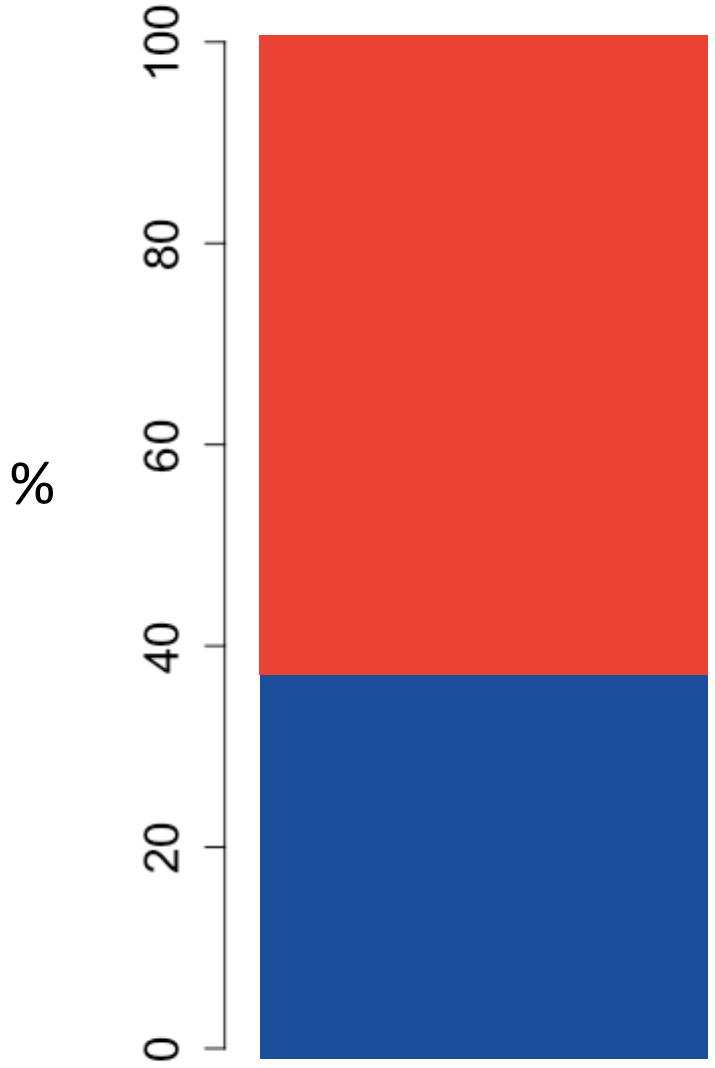
Results – Reference to CONSORT (N=73)



N=46 (63%) with **no** CONSORT checklist

N= 27 (37%) with **some** CONSORT checklist

Results – Reference to CONSORT (N=73)

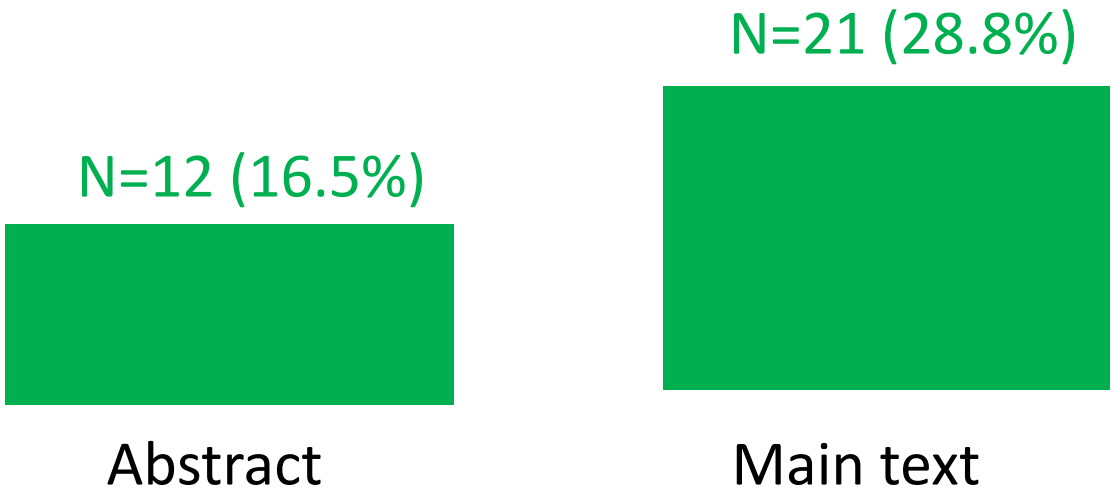


Of N=46 with **no** CONSORT checklist
9 (19.5%) reported both abs & rel effects

Of N= 27 with **some** CONSORT checklist
5 (18.5%) reported both abs & rel effects

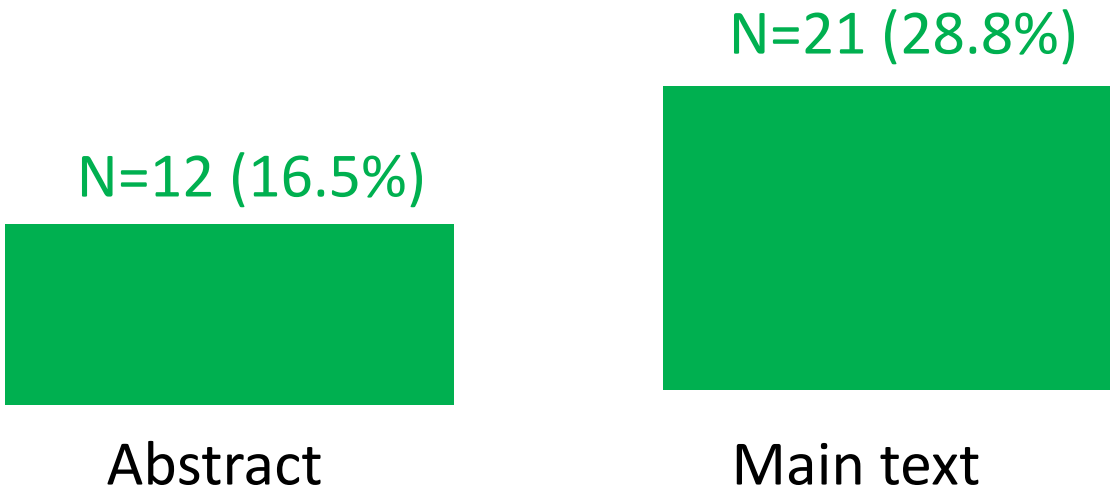
Results – Reporting of absolute effects (of total of N=73 CRTs)

Results – Reporting of absolute effects (of total of N=73 CRTs)



Absolute effects

Results – Reporting of absolute effects (of total of N=73 CRTs)



Absolute effects

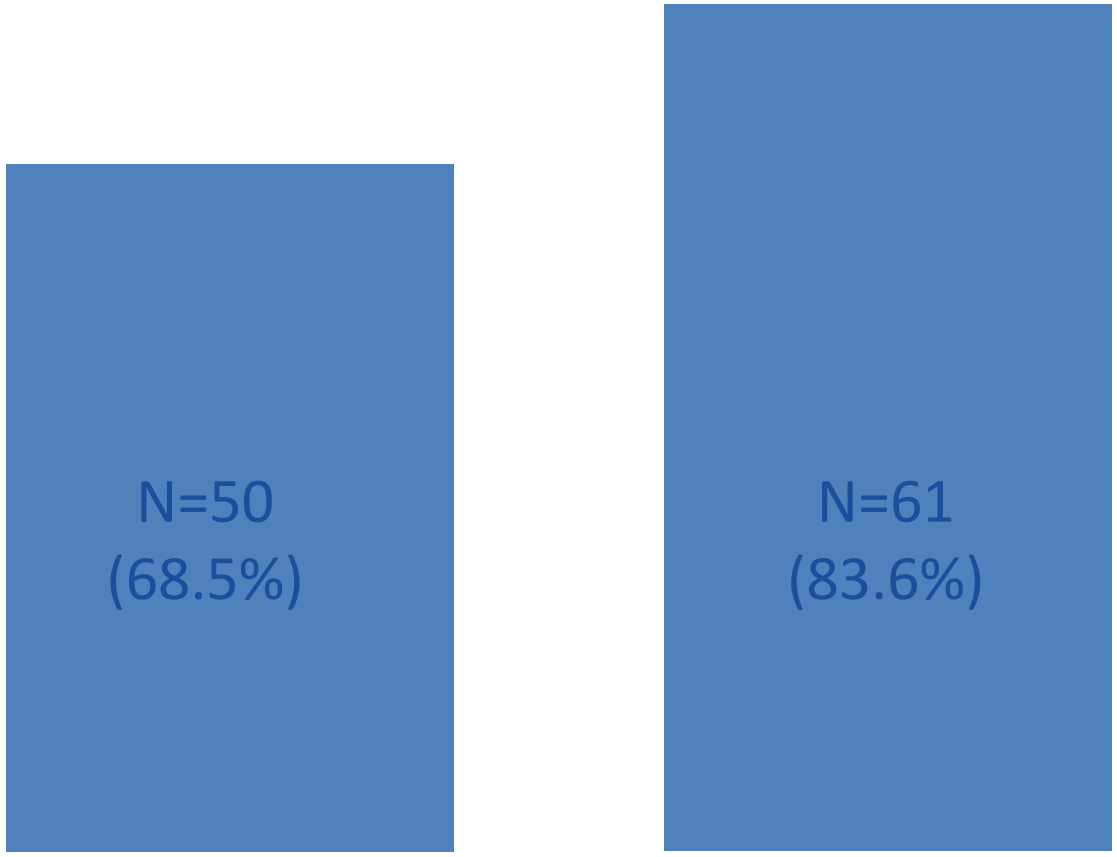
Type*	Abstract (N=12)	Main text (N=21)
Risk difference	9 (75%)	17 (81%)
NNT	2 (17%)	3 (14%)
Other#	2 (17%)	2 (10%)

N (% of 12 and 21 for abstract and main text, respectively);

* Not mutually exclusive; # Reported as difference in differences (baseline to follow-up)

Results – Reporting of relative effects (of total of N=73 CRTs)

Results – Reporting of relative effects (of total of N=73 CRTs)



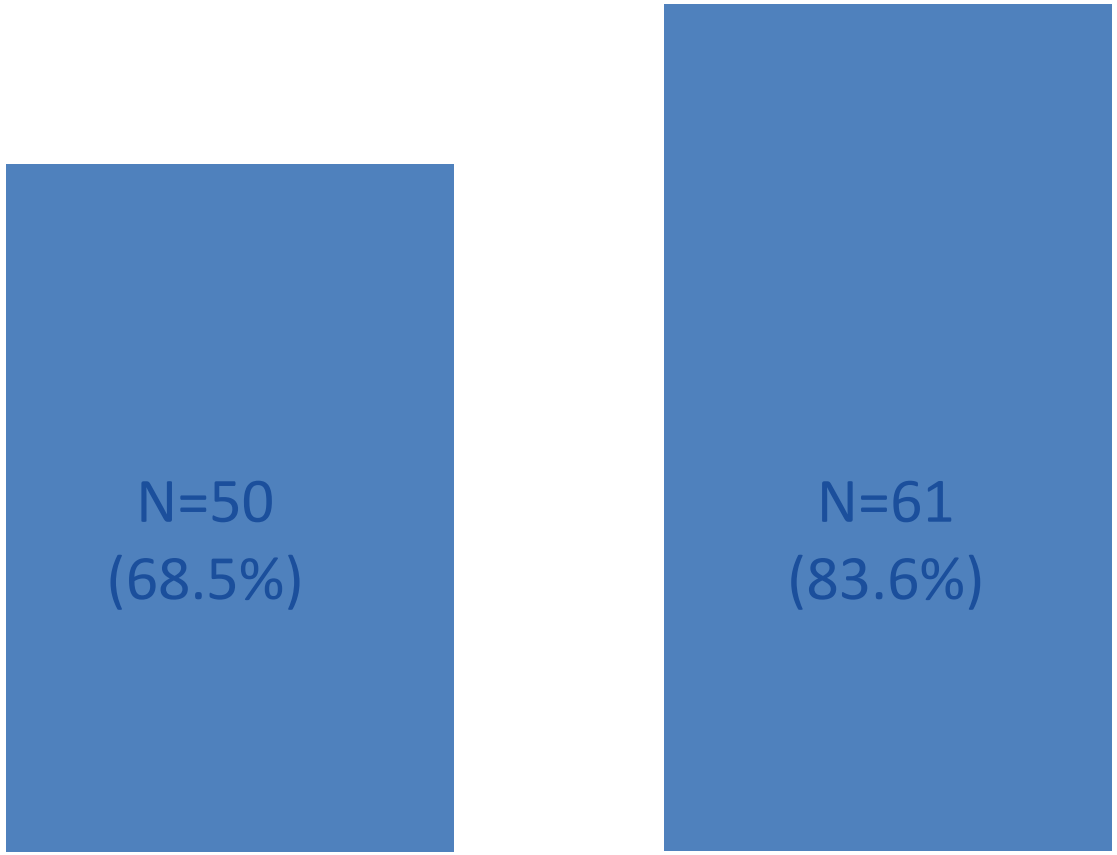
Results – Reporting of relative effects (of total of N=73 CRTs)

Type*	Abstract (N=50)	Main text (N=61)
Odds ratio	32 (64%)	39 (64%)
Risk ratio	16 (32%)	18 (30%)
Other#	3 (6%)	4 (7%)

N (% of 50 and 61 for abstract and main text, respectively);

* Not mutually exclusive;

In both abstract and main text: 2 ratio of odds ratio; 1 LATE; and in main text only ratio of cumulative incidence



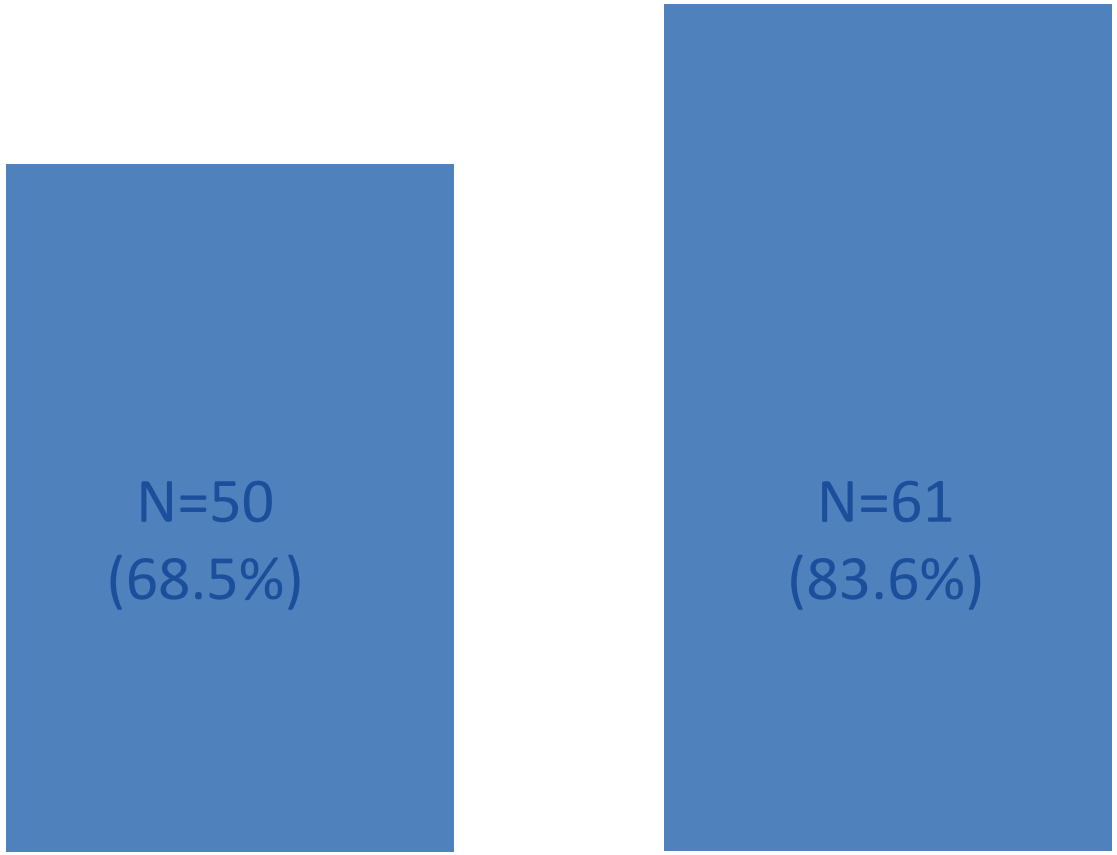
Results – Reporting of relative effects (of total of N=73 CRTs)

Type*	Abstract (N=50)	Main text (N=61)
Odds ratio	32 (64%)	39 (64%)
Risk ratio	16 (32%)	18 (30%)
Other#	3 (6%)	4 (7%)

N (% of 50 and 61 for abstract and main text, respectively);

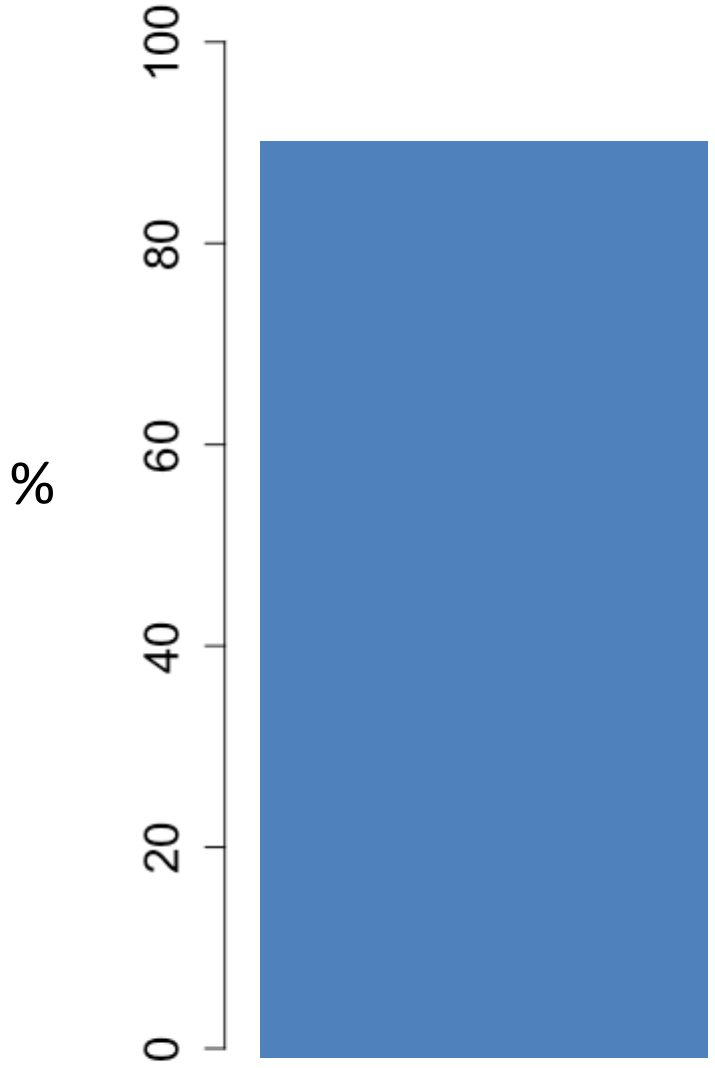
* Not mutually exclusive;

In both abstract and main text: 2 ratio of odds ratio; 1 LATE; and in main text only ratio of cumulative incidence



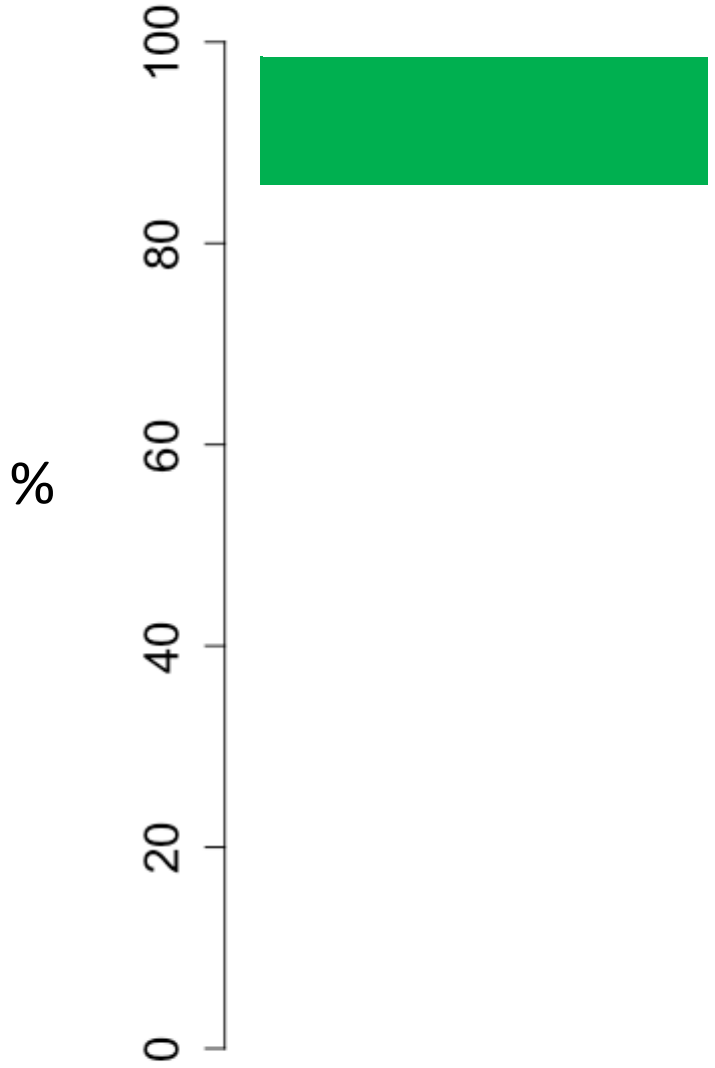
Reporting results – Analysis methods (N=73 CRTs)

Reporting results – Analysis methods (N=73 CRTs)



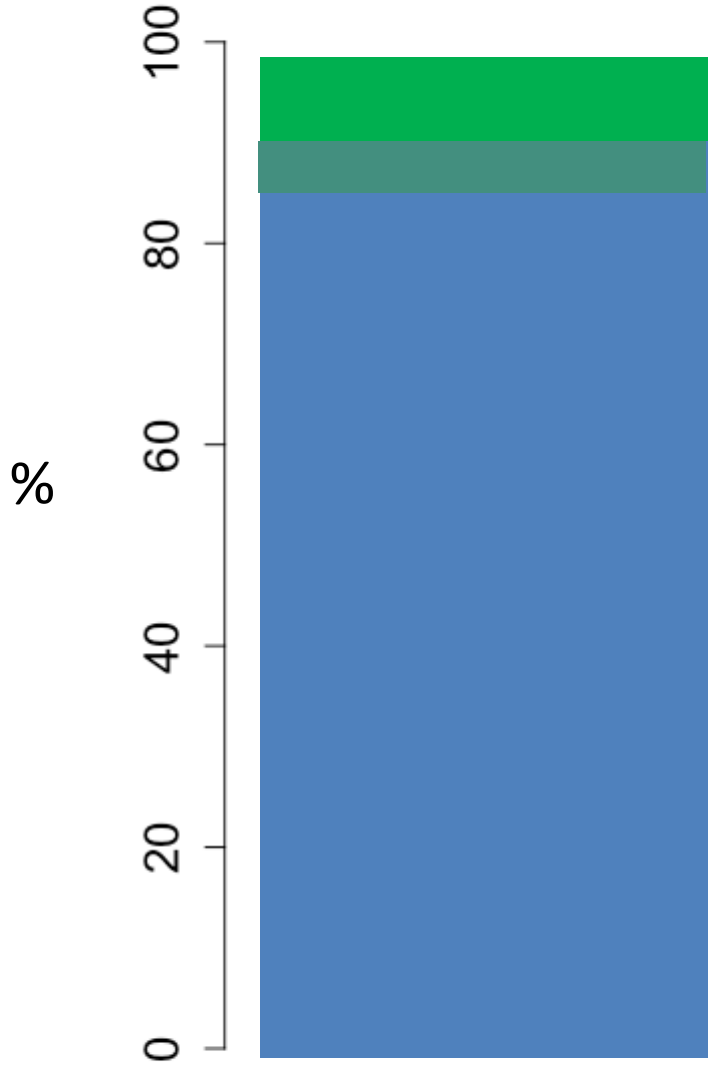
Individual-level analysis
N=64 (87.7%)

Reporting results – Analysis methods (N=73 CRTs)



Cluster-level analysis
N=11 (15.1%)

Reporting results – Analysis methods (N=73 CRTs)

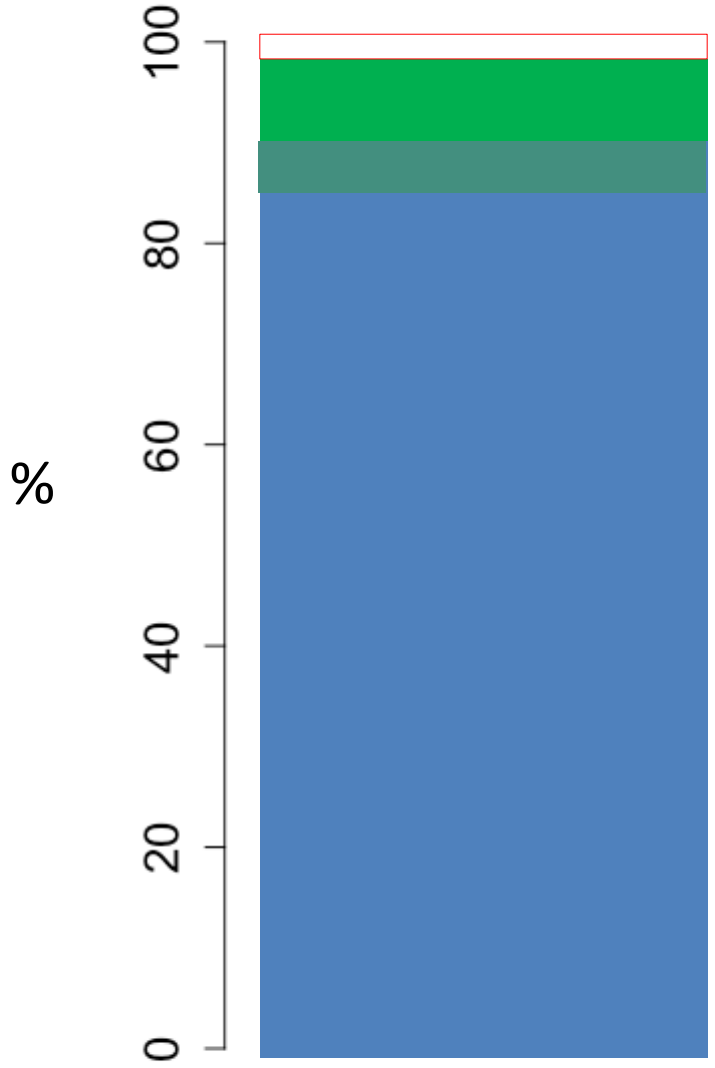


Cluster-level only: N=8 (11%)

Both analyses: N=3 (4.1%)

Individual-level only
N=61 (83.6%)

Reporting results – Analysis methods (N=73 CRTs)



Neither/Not clear: N=1

Cluster-level only: N=8 (11%)

Both analyses: N=3 (4.1%)

Individual-level only
N=61 (83.6%)

Cluster-level methods used (N=11 CRTs)

Cluster-level methods used (N=11 CRTs)

Method	N (%)
Main cluster-level summary statistic analyzed*	
Proportions	9 (82%)
Mean residuals	1 (9%)
Other	2 (18%)
Method to compare cluster-level summary statistic*	
T-test	5 (46%)
Z-test	0 (0%)
Wilcoxon Rank Sum test	1 (9%)
Permutation test	1 (9%)
Other#	4 (36%)

*Categories not mutually exclusive; # 2 regression of cluster-level proportions; 1 regression of log-cluster proportions; 1 logistic regression of dichotomized cluster-proportions

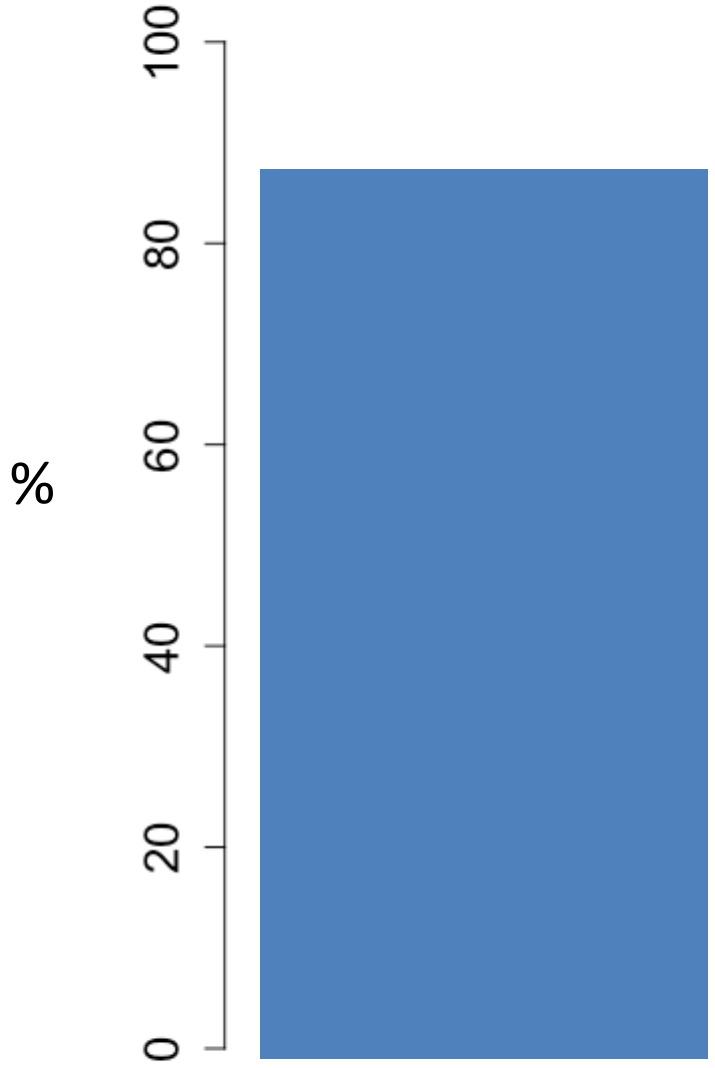
Cluster-level methods used (N=11 CRTs)

Method	N (%)
Main cluster-level summary statistic analyzed*	
Proportions	9 (82%)
Mean residuals	1 (9%)
Other	2 (18%)
Method to compare cluster-level summary statistic*	
T-test	5 (46%)
Z-test	0 (0%)
Wilcoxon Rank Sum test	1 (9%)
Permutation test	1 (9%)
Other#	4 (36%)

*Categories not mutually exclusive; # 2 regression of cluster-level proportions; 1 regression of log-cluster proportions; 1 logistic regression of dichotomized cluster-proportions

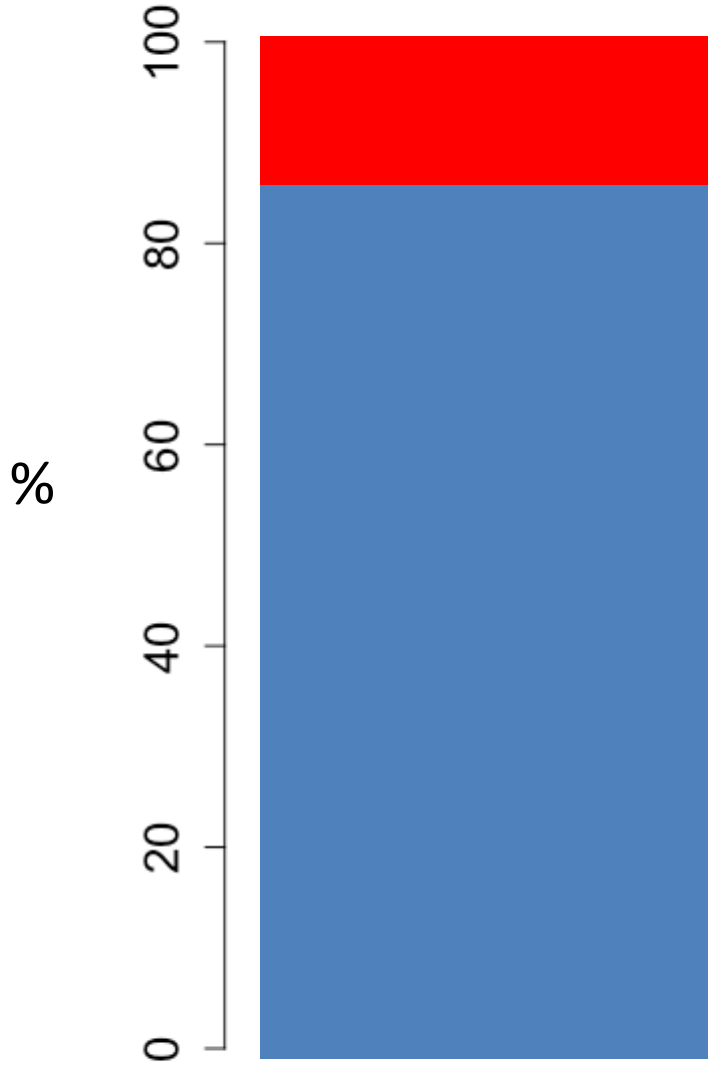
Individual-level methods used (N=64 CRTs)

Individual-level methods used (N=64 CRTs)



Accounted for clustering
N= 55 (86%)

Individual-level methods used (N=64 CRTs)

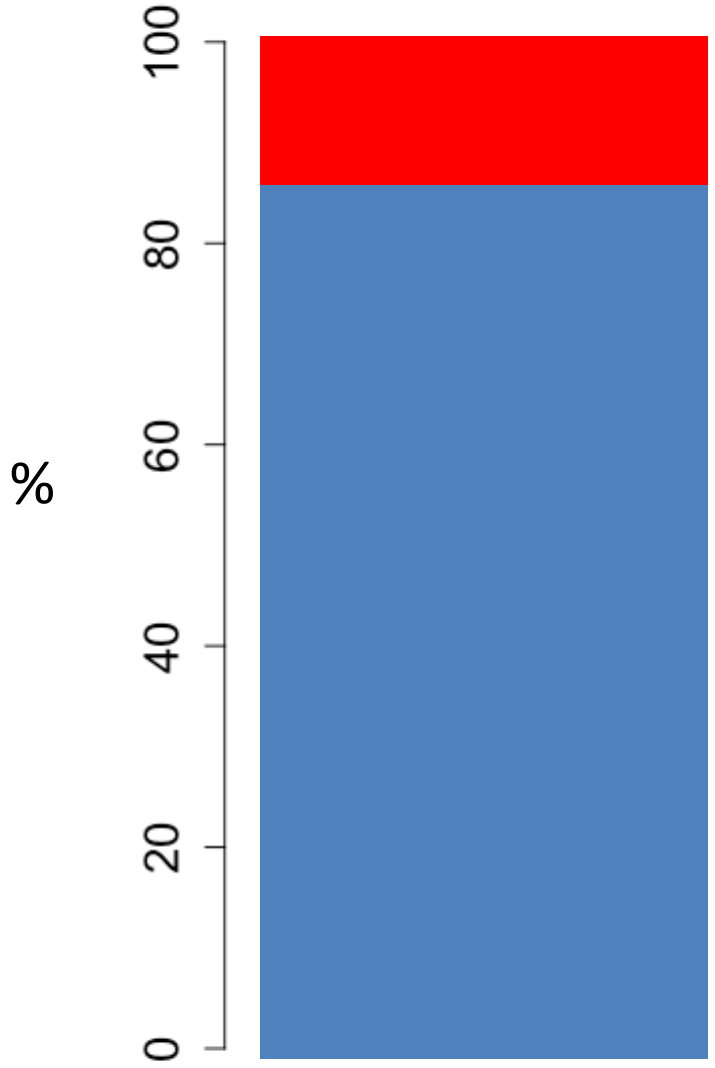


Did not account for clustering
N=9 (14%)*

Accounted for clustering
N= 55 (86%)

* Of which 1 article said no evidence of clustering and therefore did not account for it.

Individual-level methods used (N=64 CRTs)



Did not account for clustering
N=9 (14%)*

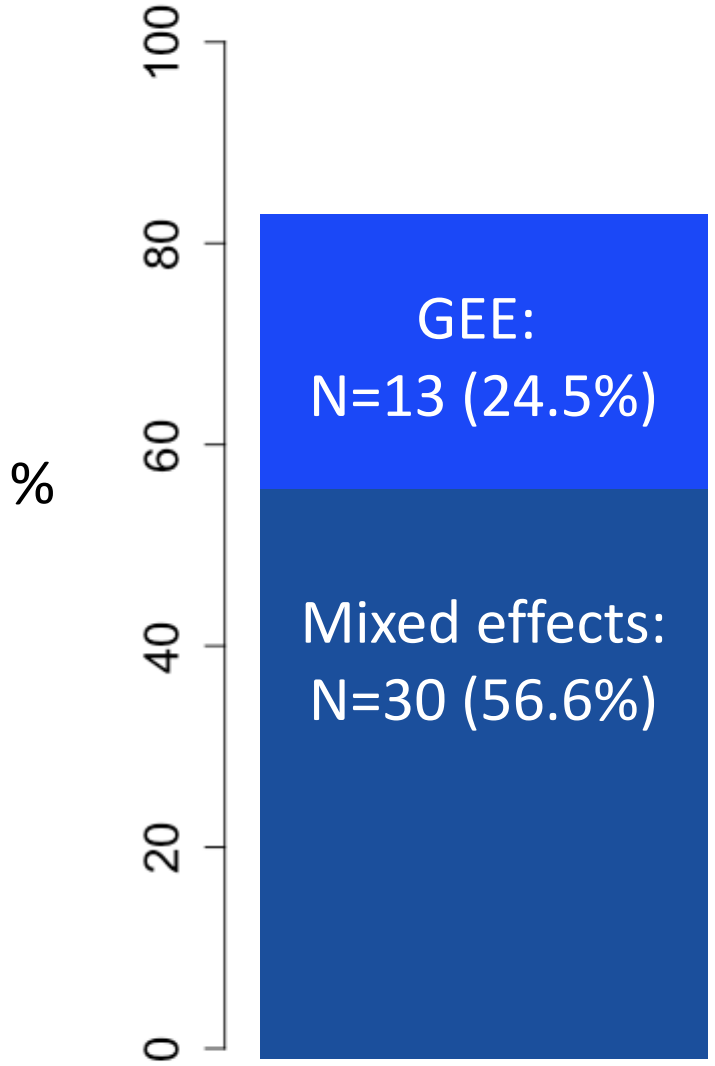
Accounted for clustering
N= 55 (86%)

N=53 regression methods
N=2 statistical test only

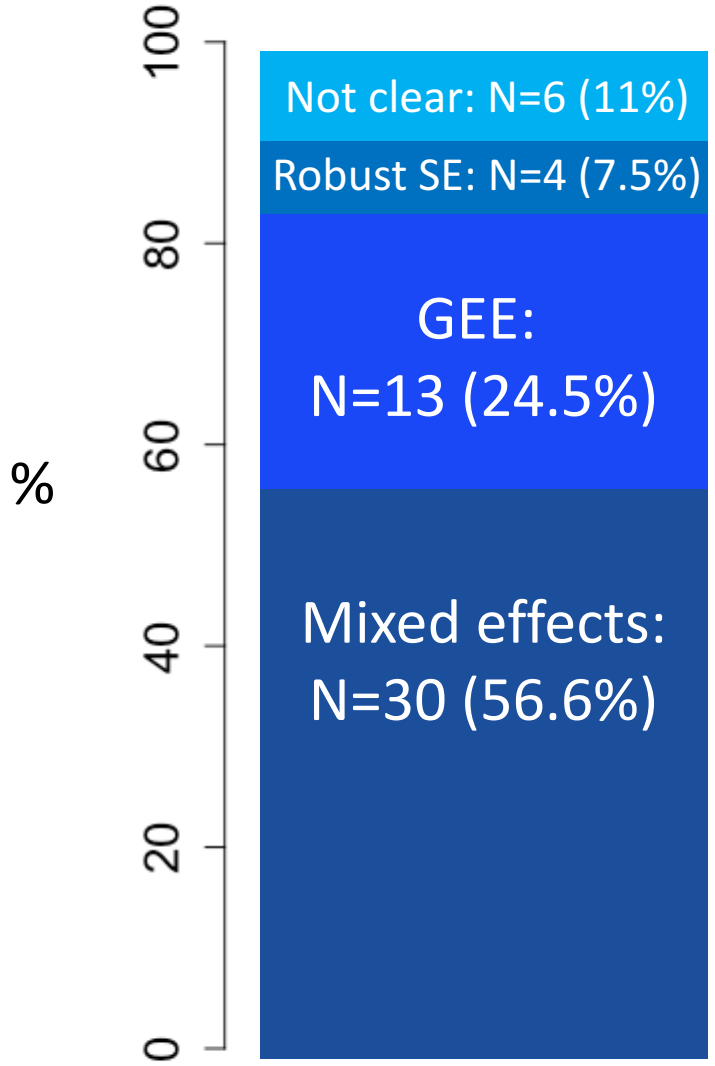
* Of which 1 article said no evidence of clustering and therefore did not account for it.

Individual-level regression methods used (N=53 CRTs)

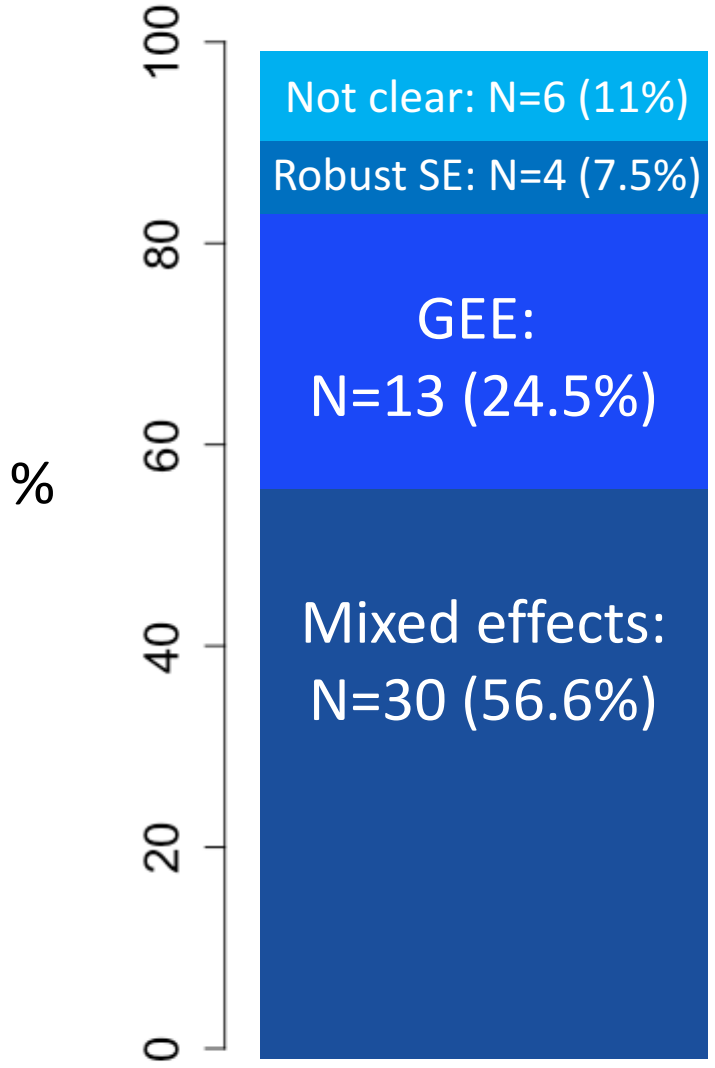
Individual-level regression methods used (N=53 CRTs)



Individual-level regression methods used (N=53 CRTs)



Individual-level regression methods used (N=53 CRTs)



- Median (IQR) # clusters:
  - 29 (20,44)
- Concern
  - Was valid inference used?
  - GEE “small sample” correction
  - DF correction for mixed model

Interlude: Shameless advertising

- Implementing “small sample” correction for GEE
  - New Stata package for xtgeebcv on SSC
  - Manuscript accepted at Stata Journal
  - Joint work with John Gallis (Duke) & Fan Li (Yale)

Interlude: Shameless advertising

- Implementing "small sample" correction for GEE
  - New Stata package for xtgeebcv on SSC
  - Manuscript accepted at Stata Journal
  - Joint work with John Gallis (Duke) & Fan Li (Yale)

- Relative effects in RCTs
  - Manuscript accepted at Annals of Global Health (joint with John Gallis)
  - Review of methods for RR and OR
  - SAS, R, Stata, SPSS code for OR and RR estimation
  - Both iRCTs and CRTs

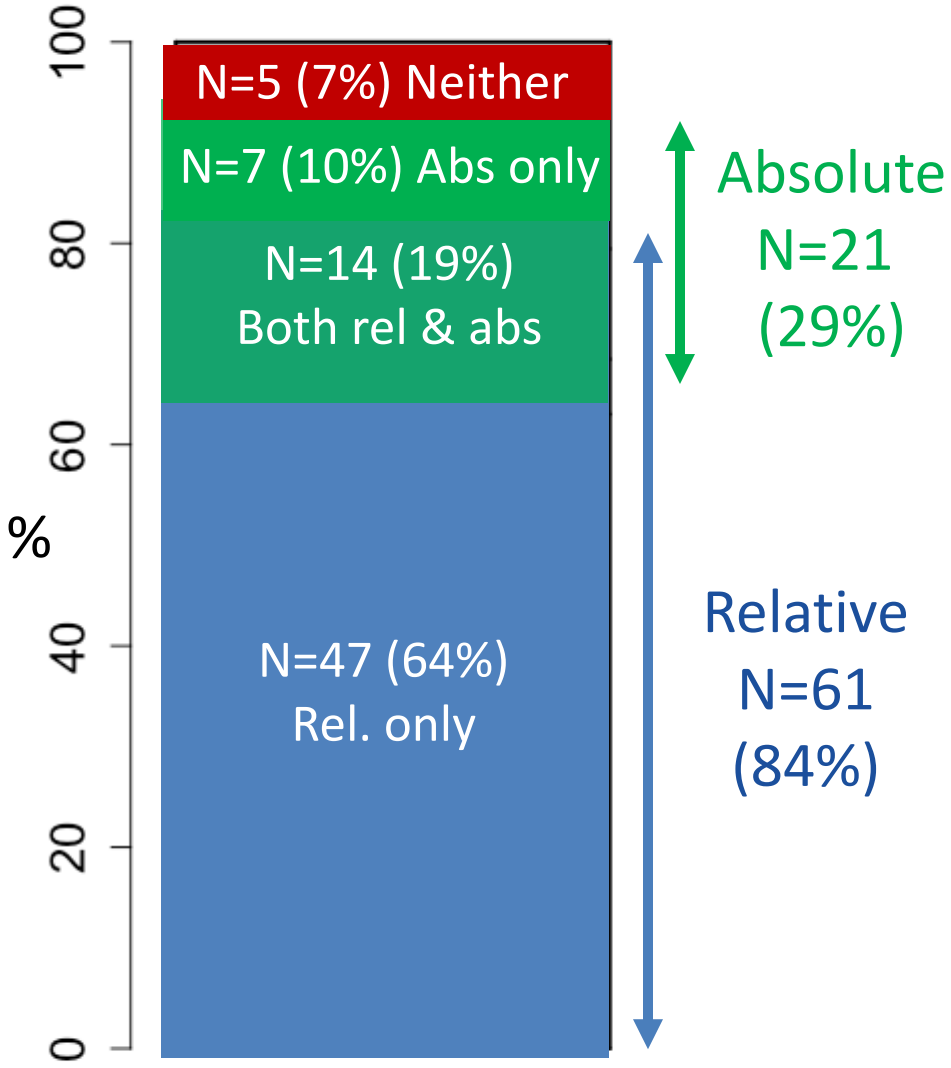
Summary

REVIEW OF REPORTING OF BINARY OUTCOMES IN CRTs

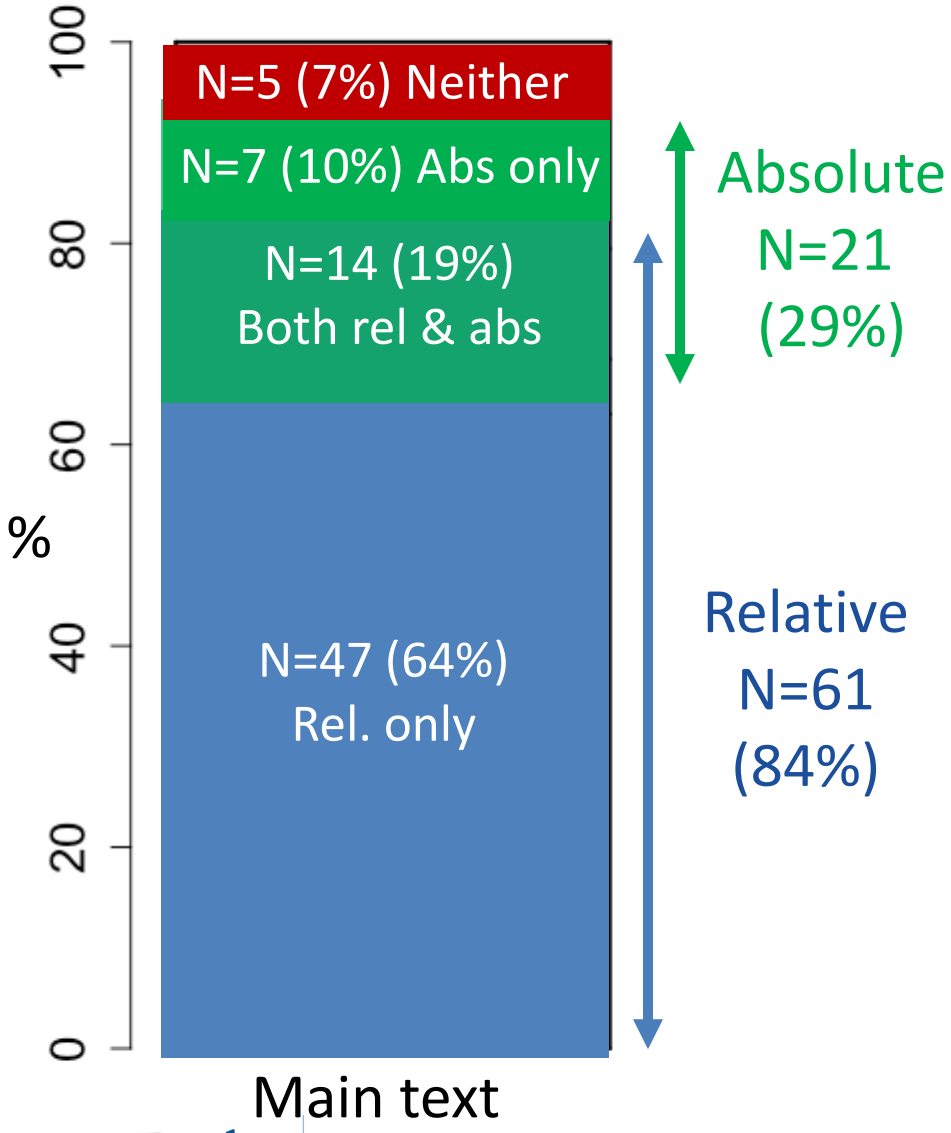
Goals of the review

1. Summarize effect measures for binary outcomes in CRTs
2. Compare to CONSORT recommendations
3. Summarize the statistical methods used
4. Identify opportunities to raise awareness of issues and to clarify methods to the community
5. Highlight the pros and cons of the “crowd-sourced” approach

Summary

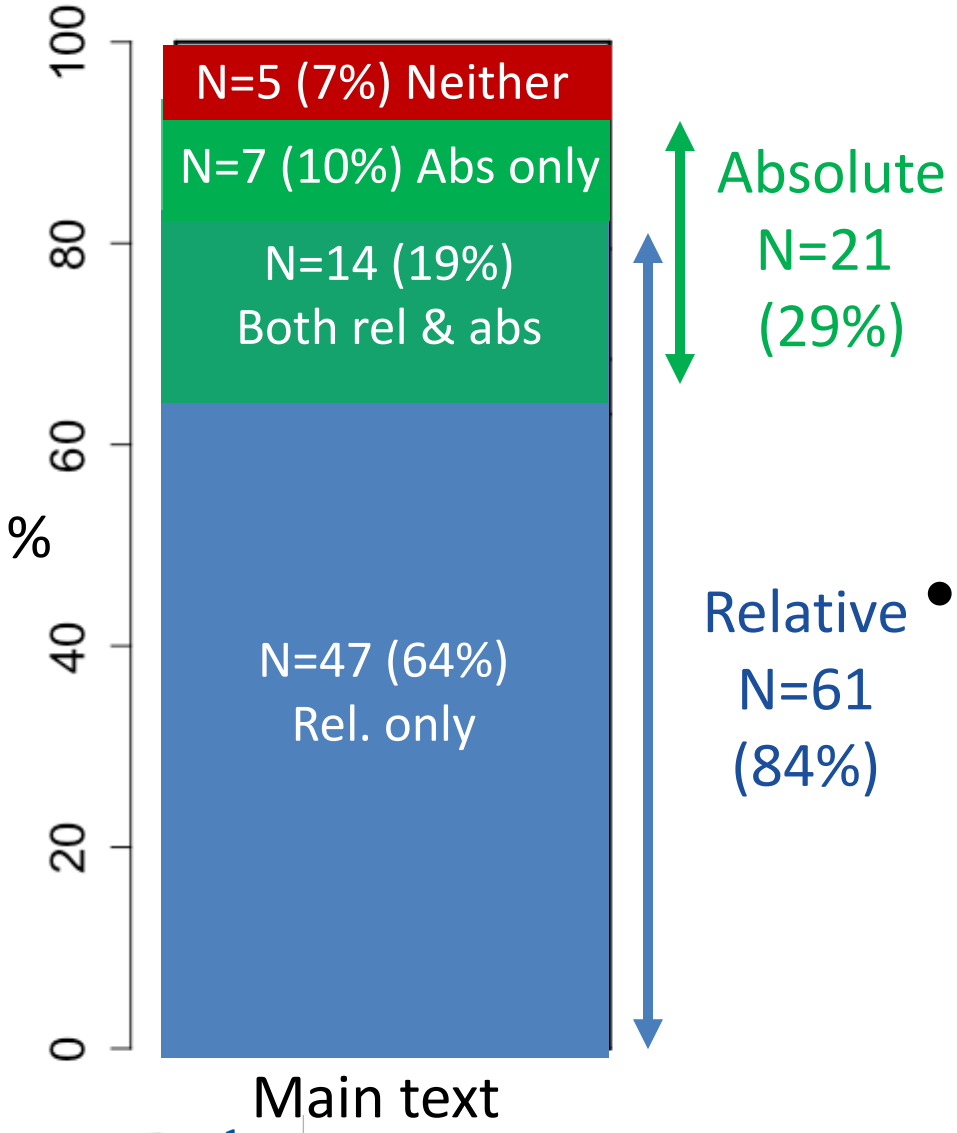

100
80
60
40
20
0
%
N=5 (7%) Neither
N=7 (10%) Abs only
N=14 (19%)
Both rel & abs
N=47 (64%)
Rel. only
Absolute
N=21
(29%)
Relative
N=61
(84%)
Main text

Summary



- Reporting (of 73 CRTs)
  - Most (63%) no CONSORT checklist
  - Most (64%) only relative effects
  - Usually odds ratio
  - Potential for misinterpretation

Summary



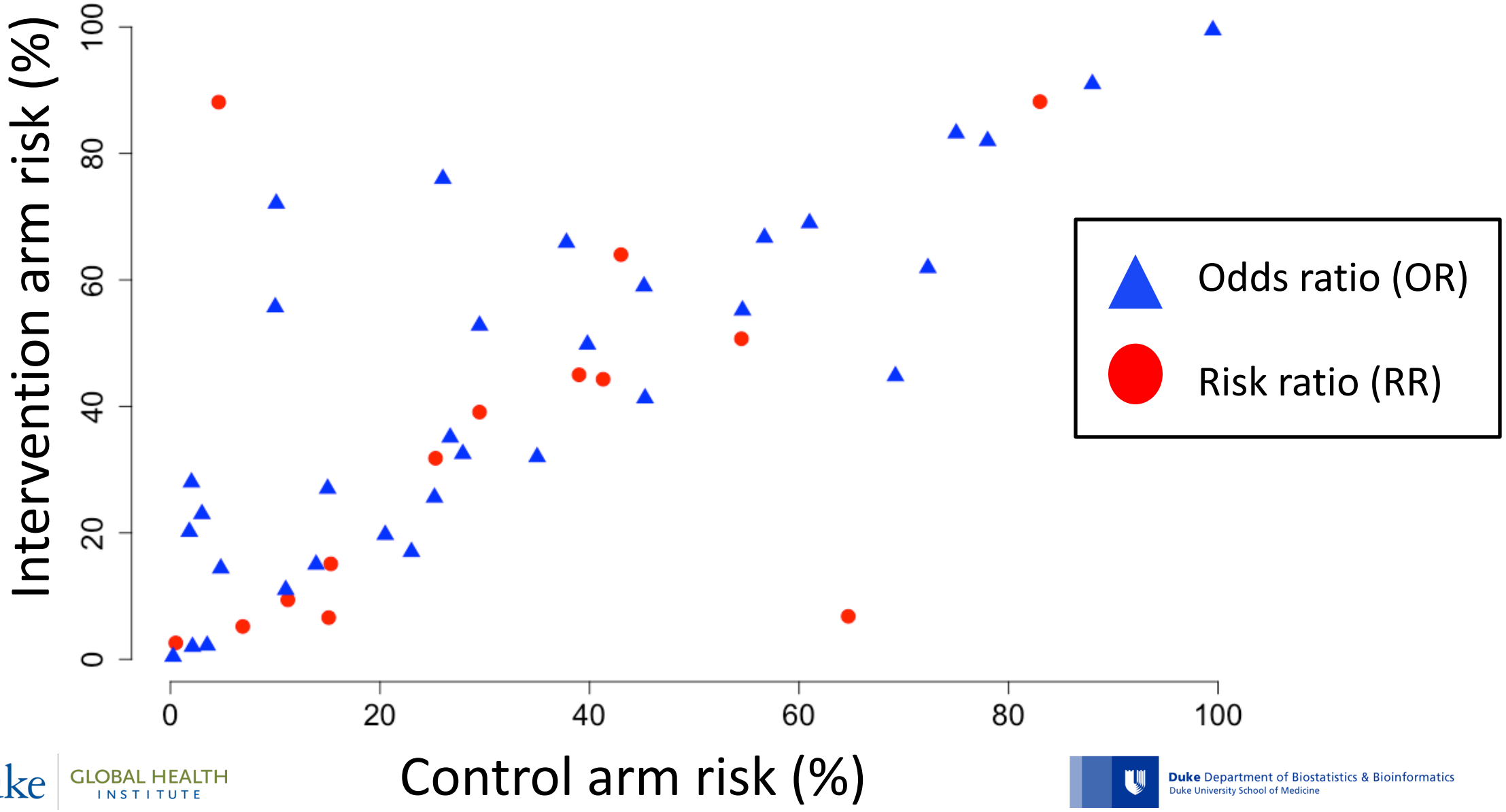
- Reporting (of 73 CRTs)
  - Most (63%) no CONSORT checklist
  - Most (64%) only relative effects
  - Usually odds ratio
  - Potential for misinterpretation
- Analysis
  - Methods for RR and RD may have model fitting issues vs. OR methods
  - Lack of awareness of methods?

Thank you!

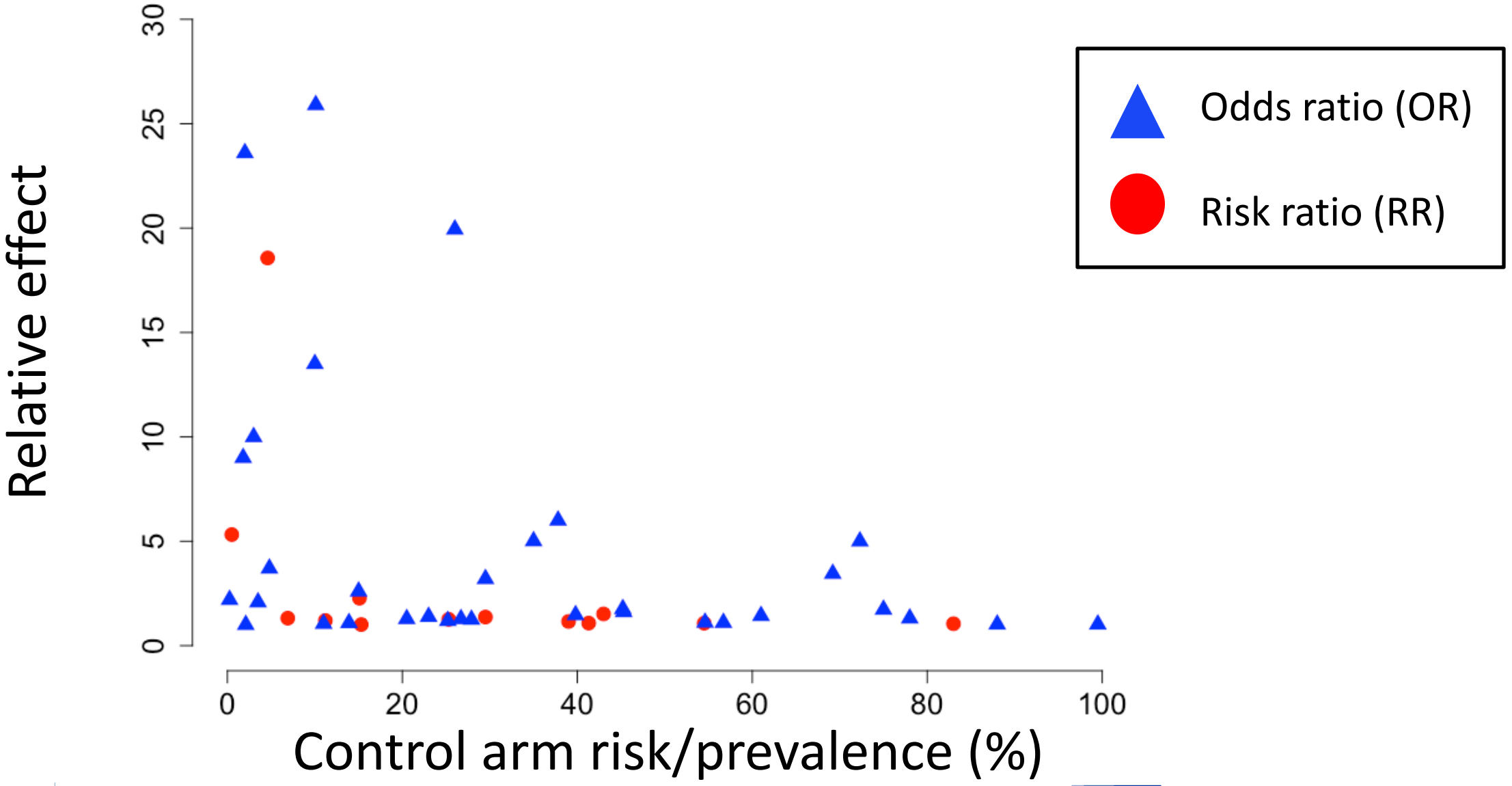
Extra slides

Results – Details - Relative effects in main text (N=61)

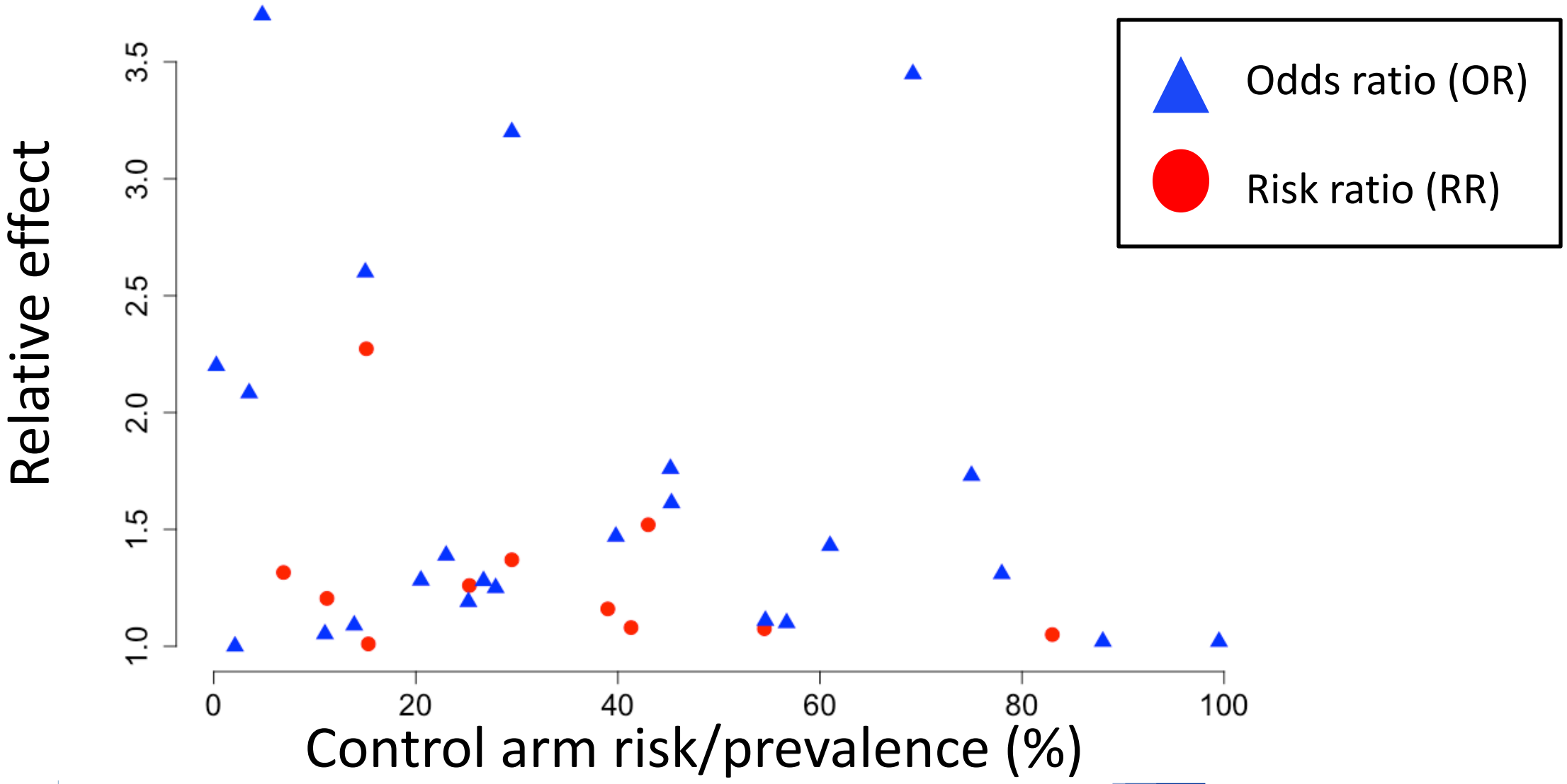
Results – Details - Relative effects in main text (N=61)



Results – Details - Relative effects in main text (N=61)



Results – Details - Relative effects in main text (N=61)



Results – 11 Journals with >1 CRT (out of N=73 CRTs in 48 journals)

Journal Name	CONSORT 2010 Statement	CONSORT 2010 Extension to Cluster Trials	Count
The Lancet Global Health	Explicitly Required	Explicitly Required	4
The Lancet			5
Cancer		Implicitly Required - EQUATOR	3
Plos One			8
Plos Medicine			2
BMC Health Services Research	Recommended	Implicitly Recommended - EQUATOR	2
Bulletin of WHO			2
BMC Public Health		No mention	2
BMC Medicine			2
Journal of Adolescent Health	No mention		3
JAIDS			3
Total Papers			34/73 (47%)
Total Journals			11/48 (23%)

Background

- Expect relative effects mostly reported as odds ratios
- Potential for misinterpretation if outcome is common and OR interpreted as a risk ratio
- Methods for RR and RD less well known as those for OR
- What is done in practice?

Overview

Last year's conference + 2 other workshops:

Undertook a crowd-sourced review of reporting of binary outcomes in CRTs.

↓

Today:

Report back on results, seek input & share reflections