Adaptive Platform Trial Scientific Meeting

September 28 – 29 • Toronto, Canada

CanTreatCOVID

Canadian Adaptive Platform Trial of Treatments for COVID in Community Settings

P-value and Bayesian analysis in randomized-controlled trials in child health research published over ten years, 2007 to 2017

Alex Aregbesola, MD, PhD
Endowered Chair, Pediatric Emergency Medicine
Assistant Professor, Department of Pediatrics and Child Health
The Children's Hospital Research Institute of Manitoba
Rady Faculty of Health Sciences, University of Manitoba

Collaboration / Affiliation

The Children's Hospital Research Institute of Manitoba, Winnipeg, MB, Canada

Department of Pediatrics and Child Health, Max Rady College of Medicine, Rady Faculty of Health Sciences, University of Manitoba, Winnipeg, MB, Canada

Department of Pediatrics and the Alberta Research Centre for Health Evidence (ARCHE), University of Alberta, Edmonton, Canada

Department of Pediatrics, Division of Infectious Diseases, Stanford University School of Medicine and Meta Research Innovation Center at Stanford (METRICS), Stanford, CA

The Hospital for Sick Children and the University of Toronto, Toronto, ON, Canada and University College London, London, UK

Authors

- Alex Aregbesola
- Allison Gates
- Amanda Coyle
- Shannon Sim
- Ben Vandermeer
- Megan Skakum
- Despina Contopoulos-Ioannidis
- Anna Heath
- Lisa Hartling
- Terry P. Klassen

BACKGROUND/RATIONALE

Overview of two statistical frameworks – Bayesian statistics versus null hypothesis significance testing

Limitations of the frequentist approach in randomized-controlled trials

BACKGROUND/ RATIONALE

Original Investigation

Evolution of Reporting *P* Values in the Biomedical Literature, 1990-2015

David Chavalarias, PhD; Joshua David Wallach, BA; Alvin Ho Ting Li, BHSc; John P. A. Ioannidis, MD, DSc

IMPORTANCE The use and misuse of *P* values has generated extensive debates.

OBJECTIVE To evaluate in large scale the *P* values reported in the abstracts and full text of biomedical research articles over the past 25 years and determine how frequently statistical information is presented in ways other than *P* values.

Editorial page 1113

Supplemental content at jama.com

CME Quiz at jamanetworkcme.com

BACKGROUND/ RATIONALE

RESEARCH QUESTIONS

- Has the inferential statistical framework in child health research changed?
- Is there a clustering around P-values of significance in RCTs in child health research papers?
- Is the Bayesian method a good alternative in conducting child health research?
- What is the expected impact of our study on the future of child health research?

OBJECTIVES

- We investigated the extent, if any, to which the inferential statistical framework in child health research has changed over 10-years
- Is there a clustering around P-values of significance in RCTs in child health research papers?

METHODS

- A review of randomly selected RCTs
- The present protocol has been registered within the Open Science Framework platform (registration ID: https://osf.io/aj2df)
- We leveraged a pre-existing sample of child health RCTs published in 2007 (n = 300) with a comparable sample of child health RCTs published in 2017 (n=300)
- A total of 600 RCTs

METHODS

Inclusion criteria

- ✓ Published RCTs in child health
- ✓ Full-text articles in the English language.
- ✓ No restriction on settings in which the study was conducted, intervention, comparator or the type of outcome.
- ✓ Search strategy developed and executed.

Bayesian/Frequentist Data Extraction Guidelines

Data Extraction Guidelines:

***Use ALL AVAILABLE INFORMATION (including protocols, companion articles references in the publication, and records of trial registration) to complete data extraction. We will include a MAXIMUM OF THREE sources per trial:**

1. The trial identified as part of our sample;
2. The trial register, if available; and
3. EITHER the published protocol or methods document, if cited in our original study (first choice) OR the sentinel trial in the case of multiple publications, if cited in our original study (second choice).

If trial registration is not declared, search, in order, the trial title/key words, corresponding author, first author, and/or last author in each of:

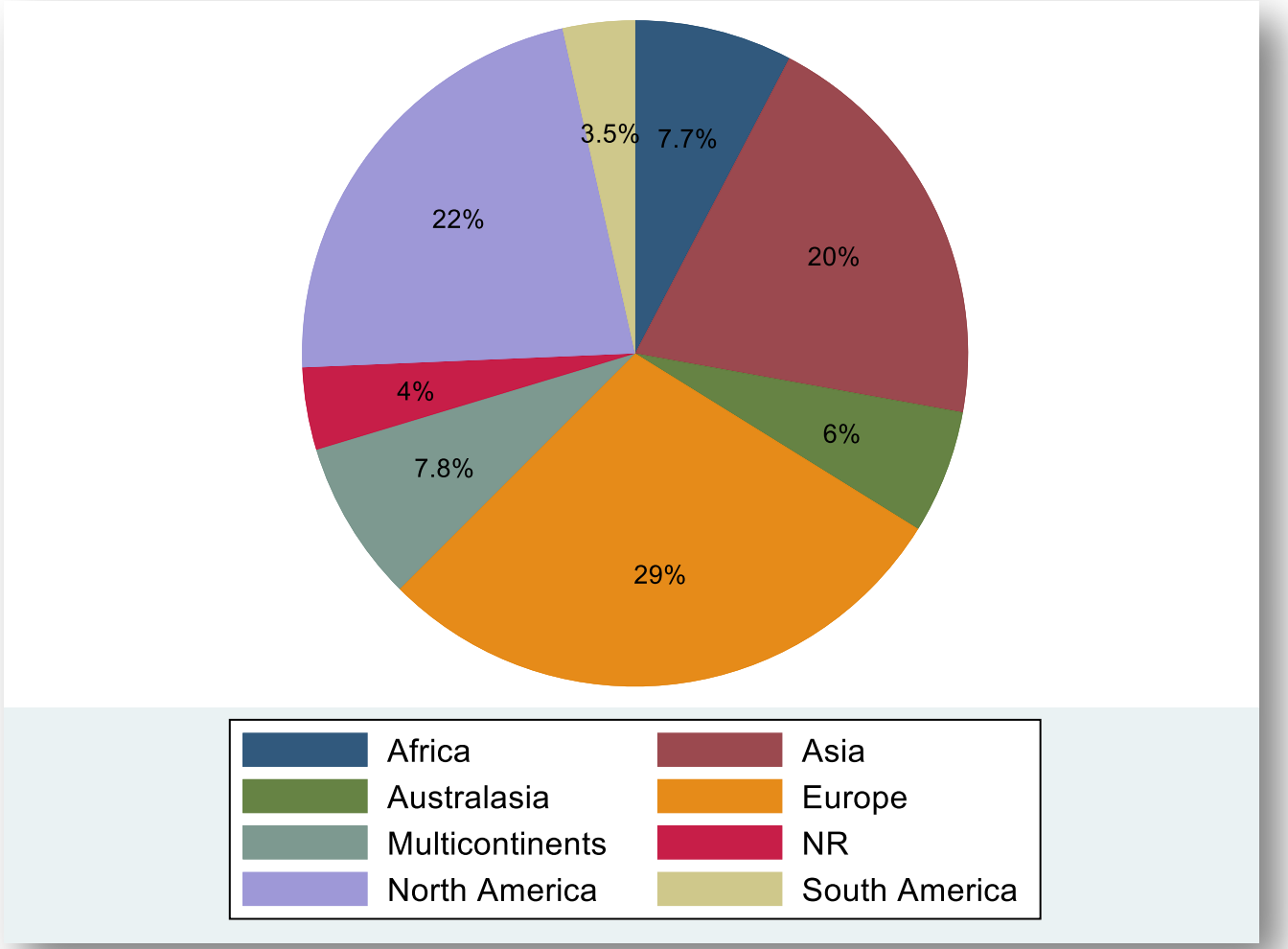
1. ICTRP (apps.who.int/trialsearch/)
2. Current Controlled Trials (www.controlled-trials.com/mrct/ -- select all 5 registers included in the meta-register)
3. Google

METHODS

DATA ANALYSIS

- We analyzed the data using
- Stata (v. 16.1; StataCorp, College Station, Texas, United States) and R.
- We used a non-informative Cauchy prior with parameters 0 and 2.5 (as recommended by Gelman)
- We also used “Jeffreys non-informative prior” for the Beta distribution and a Dirichlet with all parameters set to 0.5.

RESULTS: RCTs characteristics
Distribution of RCTs by continent



Distribution of RCTs by continent and year of publication

Continent	Publication year		Total
	2007	2017	
Africa	18	28	46
Asia	39	**82**	121
Australasia	17	19	36
Europe	**105**	67	**172**
Multi-continent	23	24	47
NR	17	**7**	24
North America	71	62	133
South America	**10**	11	**21**
Total	300	300	600
P for difference (Pearson X^2) = 0.000			

Distribution of RCTs by continent and year of publication

Continent	Mean difference	95% CI lower	95% CI upper	Probability difference >=0
Africa	0.033	-0.010	0.076	**0.937**
Asia	**0.141**	0.077	0.204	**1.000**
Australasia	0.007	-0.033	0.043	0.634
Europe	-0.125	-0.196	-0.054	0.000
Multi-continent	0.003	-0.040	0.046	0.560
North America	-0.030	-0.096	0.036	0.188
NR	-0.033	-0.066	-0.002	0.018
South America	0.003	-0.028	0.033	0.580

CI – Credible interval
NR – Not reported

Distribution of RCTs type by year of publication

RCT type	Publication year		Total
	2007	2017	
Cluster	20	38	**58**
Crossover	21	12	33
Factorial	7	1	8
Other (specify)	5	2	7
Parallel	244	243	**487**
Split body	3	4	7
	300	300	600
P for difference(Fisher's exact) = 0.013			

Distribution of RCTs type by year of publication

RCT type	Mean difference	95% CI lower	95% CI upper	Probability difference >=0
Cluster	**0.060**	0.015	0.108	**0.994**
Crossover	-0.030	-0.067	0.007	0.054
Factorial	-0.020	-0.040	-0.002	0.013
Other (specify)	-0.010	-0.030	0.007	0.123
Parallel	-0.003	-0.066	0.060	0.466
Split body	0.003	-0.014	0.022	0.648

CI – Credible interval

Distribution of RCT control type by year of publication

Control type	Publication year 2007	2017	Total
Active Intervention	138	127	**265**
No intervention	32	45	77
Other (specify)	46	53	99
Placebo	72	55	**127**
Usual care	1	0	1
Wait-list control	11	20	31
Total	300	300	600

P for difference(Fisher's exact) = 0.088

Distribution of RCTs control type by year of publication

Control type	Mean difference	95% CI lower	95% CI upper	Probability difference >=0
Active Intervention	-0.036	-0.118	0. 043	0.189
No intervention	**0.043**	-0.010	0.097	**0.942**
Other (specify)	0.023	-0.038	0.081	0.775
Placebo	-0.056	-0.122	0.009	0.045
Usual care	-0.003	-0.014	0.005	0.182
Wait-list control	**0.030**	-0.004	0.066	**0.950**
CI – Credible interval				

Distribution of RCTs number of centers by year of publication

Number of centers	Publication year 2007	2017	Total
Multicenter	112	132	244
Single center	142	167	**309**
Unclear	46	1	47
Total	300	300	600
P for difference(Fisher's exact)<0.001			

Distribution of RCTs hypothesis type by year of publication

RCT hypothesis	Publication year		
	2007	2017	Total
A priori/alternative	33	37	70
N/A	180	155	**335**
Null hypothesis	16	20	36
Null hypothesis + a priori	1	0	1
Other (specify)	70	88	158
Total	300	300	600
P for difference(Fisher's exact)= 0.204			

Distribution of RCTs hypothesis type by year of publication

Study hypothesis	Mean difference	95% CI lower	95% CI upper	Probability difference >=0
A priori/alternative	-0.019	-0.124	0. 085	0.361
Null hypothesis	0.004	-0.076	0.087	0.544
Null hypothesis + a	-0.009	-0.035	0.010	0.151
Other (specify)	0.024	-0.093	0.142	0.065

CI – Credible interval

Power of RCTs by year of publication

Power of trial calculated	Publication year 2007	2017	Total
No	281	274	555
Yes	19	26	**45**
Total	300	300	600
P for difference (Pearson X^2)= 0.278			

Sample size of RCTs by year of publication

Sample size calculated	Publication year 2007	2017	Total
No	176	133	309
Yes	124	167	291
Total	300	300	600
P value for difference (Fisher's exact) X^2=0.378			

Interim analysis of RCTs by year of publication

Interim analysis reported	Publication year 2007	2017	Total
No	282	288	**570**
Yes	18	12	30
Total	300	300	600
P for difference(Pearson X^2)= 0.261			

Primary outcome analysis of RCTs by year of publication

Primary outcome analysis based on	Publication year		
	2007	2017	Total
95% confidence interval	21	20	41
Other (specify)	69	78	147
P-value + Bayesian inferential	0	2	2
P-value only (frequentist)	210	200	**410**
Total	300	300	600
P for difference(Fisher's exact)= 0.487			

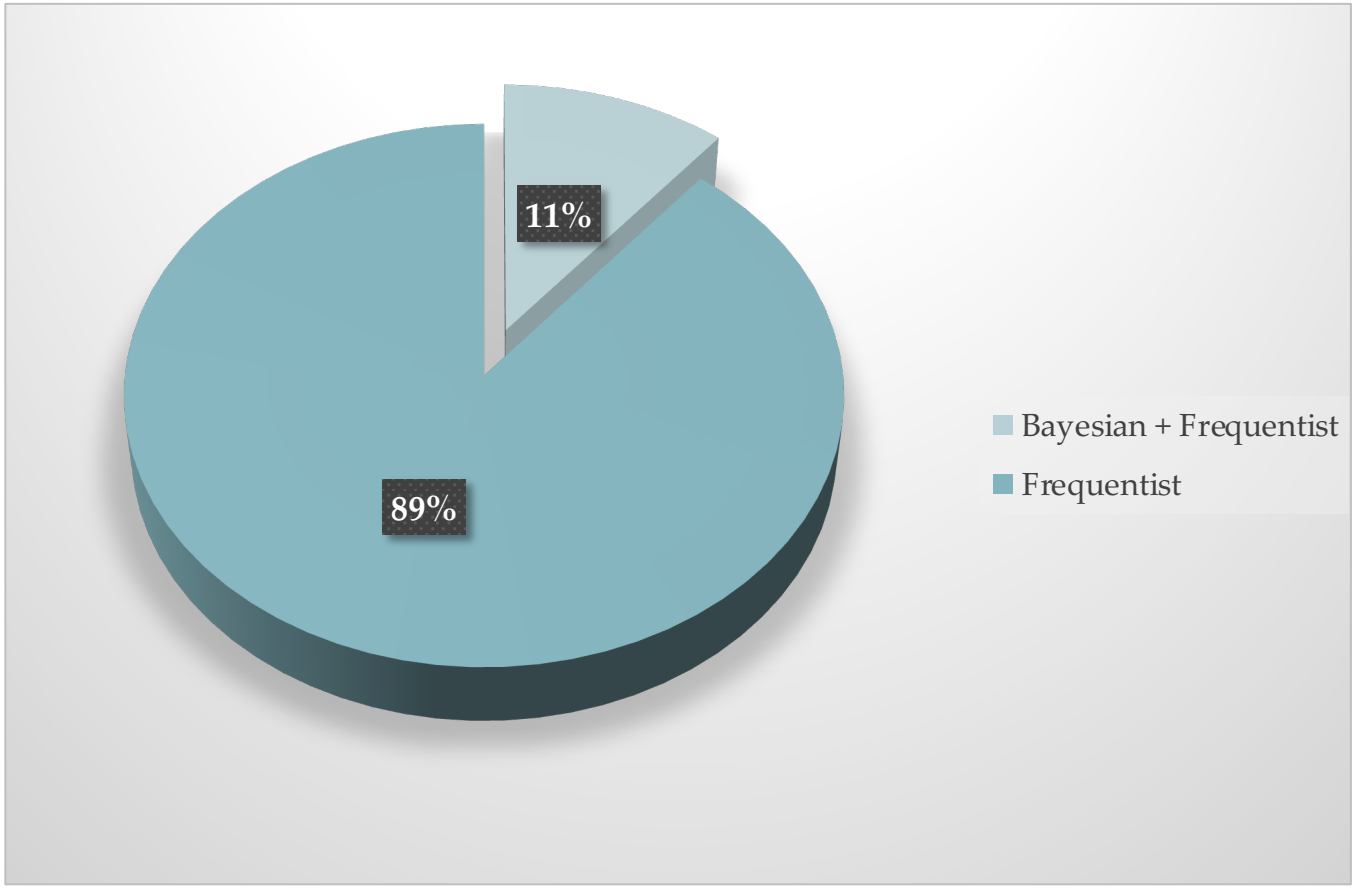
Primary outcome analysis of RCTs by year of publication

Primary outcome analysis based on	Mean difference	95% CI lower	95% CI upper	Probability difference >=0
95% confidence interval	-0.003	-0.043	0. 037	0.443
Other (specify)	0.030	-0.038	0.097	0.805
P-value + any Bayesian	**0.007**	-0.004	0.019	**0.926**
P-value only (frequentist)	-0.033	-0.109	0.039	0.186
CI – Credible interval				

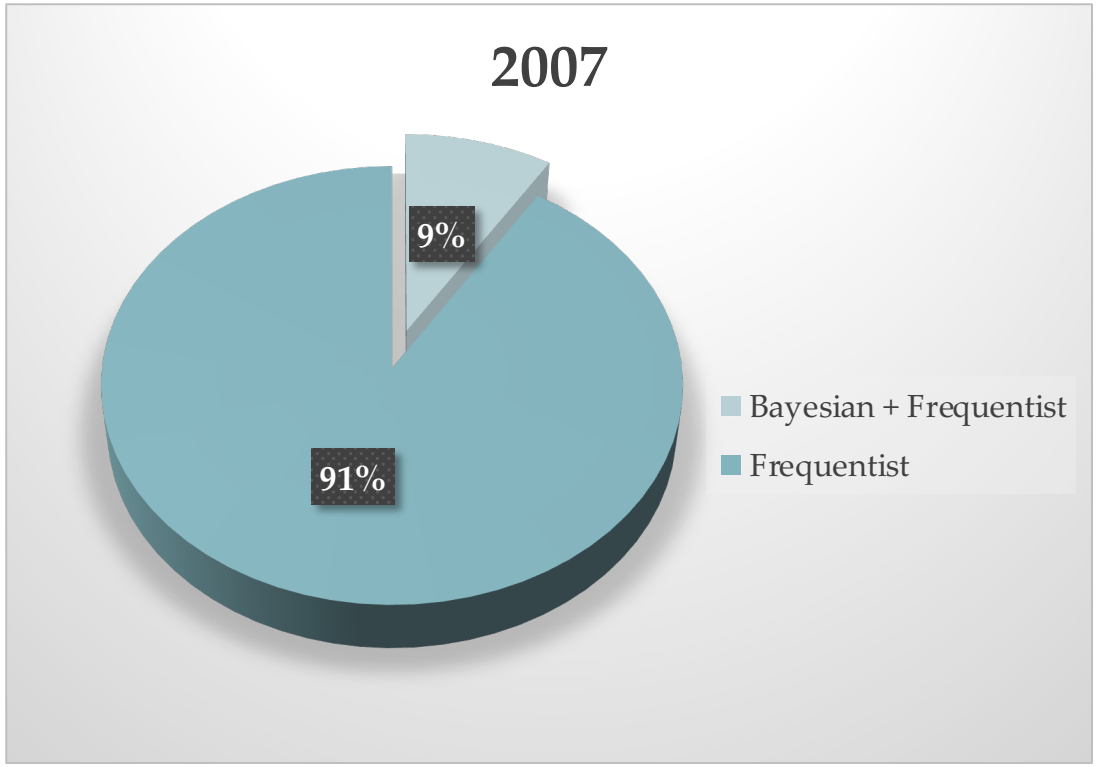
MCID of RCTs by year of publication

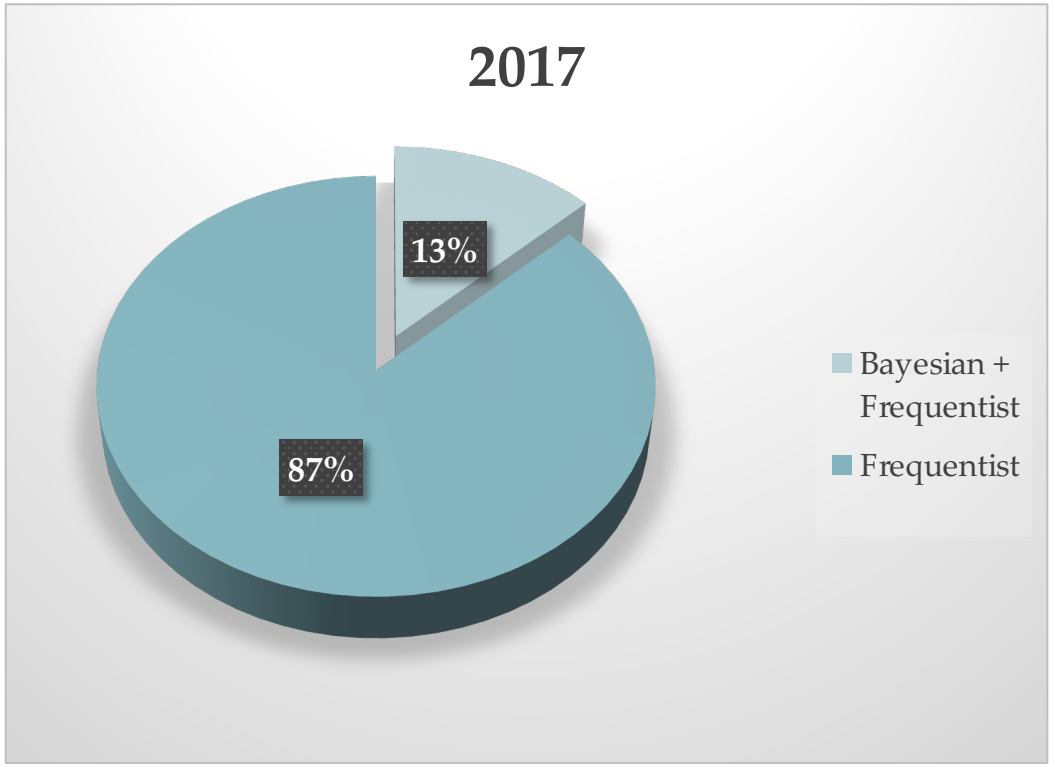
MCID reported	Publication year 2007	2017	Total
No	266	285	551
Yes	34	14	48
Total	300	299	599
P for difference(Pearson X^2)= 0.003			

Type of statistical frameworks in the 600 RCTs



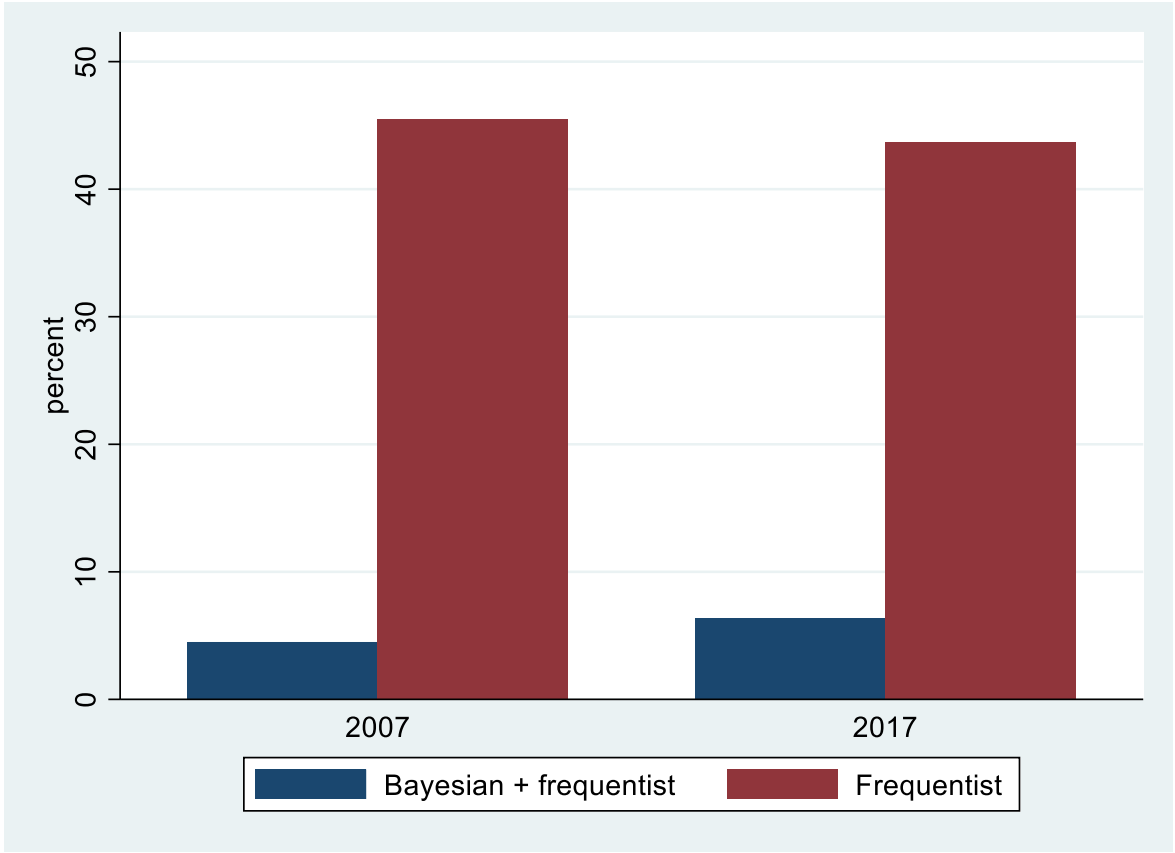
Types of statistical frameworks by year of publication in the 600 RCTs





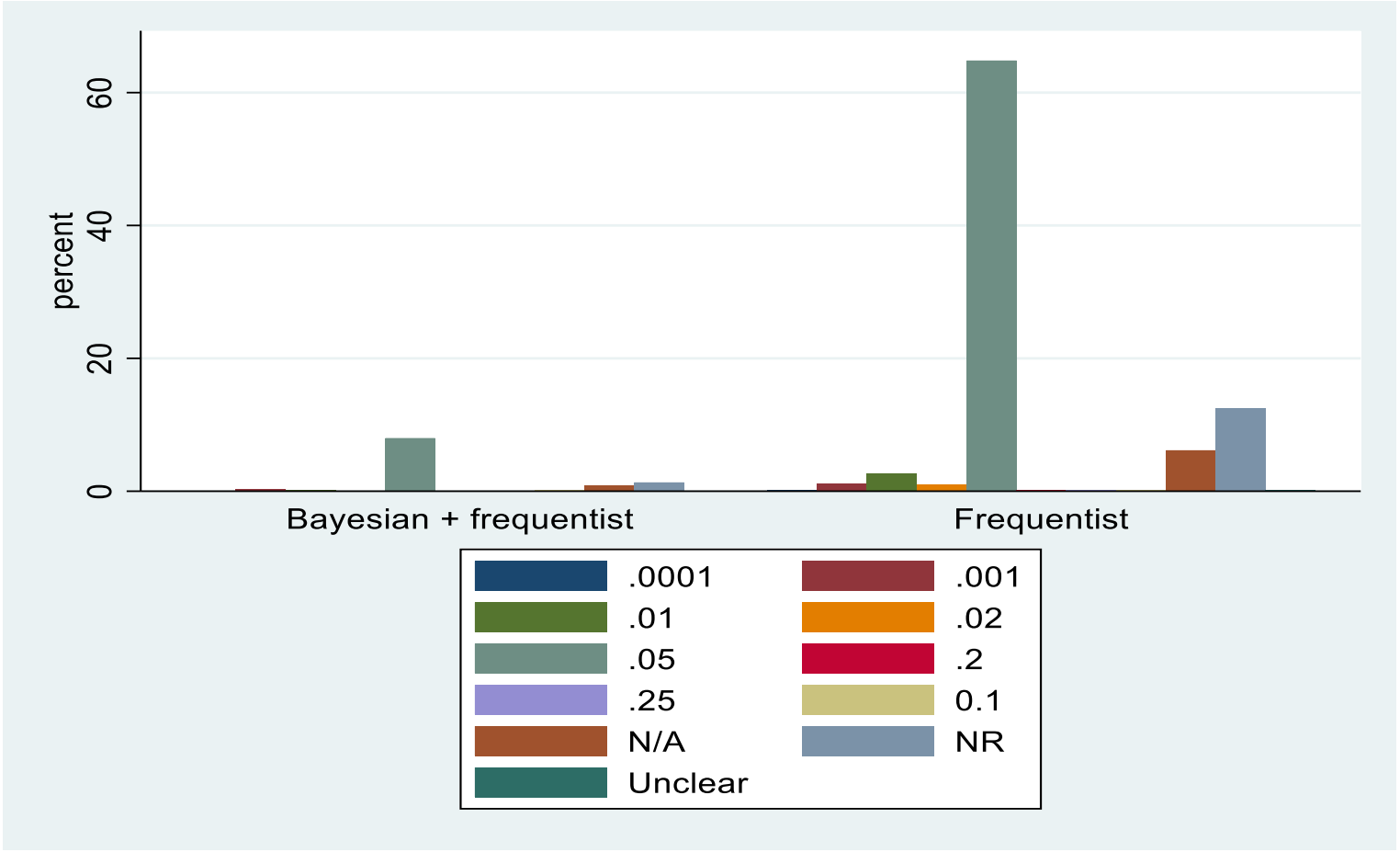
Types of statistical frameworks by year of publication in the 600 RCTs

Conclusions	Publication year 2007	2017	Total
Bayesian + Frequentist	27	38	65
Frequentist	273	262	535
Total	300	300	600
P for difference(Pearson X2)= 0.148			



Types of statistical frameworks within each year of publication, 2007 and 2017

Clustering of P-value of significance



Clustering of P-value of significance

P-value of significance	Frequency	Percent
.0001	1	0.2
.001	9	1.5
.01	17	2.8
.02	6	1.0
.05	**437**	**72.8**
.2	1	0.2
.25	1	0.2
0.1	2	0.3
N/A	42	7.0
NR	83	13.8
Unclear	1	0.2
Total	600	100.00

N/A, not applicable; NR, not reported

Clustering of P-value of significance

P-value level of significance for primary outcome	Frequentist	Bayesian plus frequentist	Total
.05	389	48	437
NR	75	8	83
N/A	37	5	42
.01	16	1	17
.001	7	2	9
.02	6	0	6
0.1	1	1	2
.0001	1	0	1
.2	1	0	1
.25	1	0	1
Unclear	1	0	1
Total	535	65	600
P for difference(Fisher's exact) = 0.659			

Association between trial characteristics and using Bayesian plus frequentist methods

Trial characteristics	Odds Ratio	95%CI lower	95%CI Upper	P-value
Sample size enrolled calculated	1.01	0.61	1.68	0.97
Power of trial calculated	0.70	0.23	2.14	0.53
Sample size calculated	1.54	0.85	2.77	0.15
Source of funding reported	1.88	0.88	4.02	0.10
Data monitoring & safety reported	0.74	0.24	2.29	0.60
95% CI reported	1.10	0.48	2.57	0.81
MCID reported	0.49	0.14	1.71	0.27
Interpretation based on 95% confidence interval	0.73	0.28	1.90	0.53
Interpretation based on effect size	0.85	0.48	1.50	0.57
Interpretation based on observed data only	0.24	0.06	0.99	0.048

Association between trial characteristics and using Bayesian plus frequentist methods

Trial characteristics	Odds Ratio	95%CI lower	95%CI Upper	P (OR>=1)
Sample size enrolled calculated	1.15	0.63	1.75	0.65
Power of trial calculated	0.10	0.21	2.14	0.33
Sample size calculated	1.60	0.72	2.77	0.93
Source of funding reported	2.07	0.80	4.02	0.96
Data monitoring & safety reported	0.98	0.21	2.29	0.38
95% CI reported	1.08	0.31	2.57	0.49
MCID reported	0.56	0.04	1.24	0.11
Interpretation based on 95% confidence interval	0.96	0.27	1.90	0.35
Interpretation based on effect size	0.90	0.42	1.44	0.32
Interpretation based on observed data only	0.43	0.03	1.18	0.068

BAYESIAN ANALYSIS

We used a non-informative Cauchy prior with parameters 0 and 2.5 (as recommended by Gelman in https://arxiv.org/pdf/0901.4011.pdf).

We also used Beta with parametres (0.5, 0.5) – this is "Jeffreys non-informative prior" for the Beta distribution and a Dirichlet with all parameters set to 0.5 (again, this is the Jeffreys prior for the Dirichlet distribution).

Essentially the Jeffreys prior is a specific non-informative prior with some good properties.

THANKYOU!

Questions!