Meta-Analysis of Rare Events

Dankmar Böhning and Antonello Maruotti

Southampton Statistical Sciences Research Institute University of Southampton

5th of September 2014

Contents

- What are rare events data?
- What are the problems with rare events trials?
- Effect estimation using Mantel-Haenszel
- Poisson with fixed and random effects
- Sensitivity analysis: the effect of excluding zero-studies

- **Zero-inflation models**
- Logistic regression modelling
- Conditional logistic regression modelling

Contents

What are rare events data?

What are the problems with rare events trials?

Effect estimation using Mantel-Haenszel

Poisson with fixed and random effects

Sensitivity analysis: the effect of excluding zero-studies

<ロト<団ト<臣ト<臣ト 3/68

Zero-inflation models

Logistic regression modelling

Conditional logistic regression modelling

motivation

- recent debate on the safety of the diabetes drug rosiglitazone
- meta-analysis (MA) by Nissen and Wolski (2007, 2010)
- number of papers Shuster *et al.* (2007), Tian *et al.* (2009), Dahabreh (2008), Friedrich *et al.* (2009), Mannucci *et al.* (2009), Kaul and Diamond (2011)

 Böhning, Mylona, Kimber (2014) focus on existing methodology to adapt to MA of rare event trials

Table: Study data of meta-analysis on rare events in Rosiglitazone and control arm; MI refers to the myocardial infarction deaths, CV to cardiovascular deaths, n is the size of the respective study arm and 'duration' refers to the study period at risk (in weeks)

		treatment arm		con	trol ar	m		
ID	study label	п	MI	CV	п	MI	CV	duration
1	49653/011	357	2	1	176	0	0	24
2	49653/020	391	2	0	207	1	0	52
3	49653/024	774	1	0	185	1	0	26
4	49653/093	213	0	0	109	1	0	26
•••			•••	•••	• • •	• • •	•••	• • •
53	49653/452	26	0	0	24	0	0	24
54	DREAM	2635	15	12	2634	9	10	156
55	ADOPT19	1456	27	2	2895	41	5	208
56	RECORD	2220	64	60	2227	56	71	_260

a second example

 NiËl-Weise *et al.* (2007) did a MA on the effect of anti-infective-treated central venous catheters on catheter-related bloodstream infection (CRBSI) in the acute care setting

- meta-analysis involved 18 clinical trials
- control group is standard catheter

Table: Meta–analysis on rare evidence data on the effect of anti-infective-treated catheter in compariosn to standard catheter; CRBSI refers to catheter-related bloodstream infection events, *n* is the size of the respective study arm

	control	arm	treatmer	nt arm
study ID	CRBSI	n	CRBSI	n
1	3	117	0	116
2	3	35	1	44
3	9	195	2	208
4	7	136	0	130
5	6	157	5	151
6	4	139	1	98
7	3	177	1	174
8	2	39	1	74

	control	arm	treatme	nt arm
study ID	CRBSI	п	CRBSI	n
9	19	103	1	97
10	2	122	1	113
11	7	64	0	66
12	1	58	0	70
13	5	175	3	188
14	11	180	6	187
15	0	105	0	118
16	1	262	0	252
17	3	362	1	345
18	1	69	4	64

a definition

MA of rare events trials deals with MA of trials which includes single-zero or double-zero trials.

A **single-zero trial** is a trial in which at least one arm are has no events. A **double-zero trial** is a trial in which both arms have no events.

Meta-Analysis of Rare Events

What are the problems with rare events trials?

Contents

What are rare events data?

What are the problems with rare events trials?

Effect estimation using Mantel-Haenszel

Poisson with fixed and random effects

Sensitivity analysis: the effect of excluding zero-studies

Zero-inflation models

Logistic regression modelling

Conditional logistic regression modelling

What are the problems with rare events trials?

popular effect measures

risk difference RD: risk in treatment arm – risk in control arm estimated by (x number of events and P is person-time)

$$\widehat{RD} = x^T / P^T - x^C / P^C$$

risk ratio RR: risk in treatment arm / risk in control arm estimated by

$$\widehat{RR} = \frac{x^T / P^T}{x^C / P^C}$$

odds ratio OR: odds in treatment arm / odds in control arm estimated by

$$\widehat{OR} = \frac{x^T / (P^T - x^T)}{x^C / (P^C - x^C)} = \frac{x^T (P^C - x^C)}{x^C (P^T - x^T)}$$

What are the problems with rare events trials?

problems can occur on two levels with zero-studies

effect measure itself

no problem for the risk difference

$$\widehat{RD} = x^T / P^T - x^C / P^C$$

▶ risk ratio and odds ratio: it might be useless (0), infinite (∞), or undefined (0/0)

$$\widehat{RR} = \frac{x^T/P^T}{x^C/P^C}$$
 and $\widehat{OR} = \frac{x^T/(P^T - x^T)}{x^C/(P^C - x^C)}$

<ロト<日、<日、<日、<日、<日、<日、<日、<日、<日、<日、<日、<日、<12/68

What are the problems with rare events trials?

problems can occur on two levels with zero-studies

uncertainty assessment

risk difference

$$\operatorname{var}(\widehat{RD}) pprox rac{x^T}{P^{T^2}} + rac{x^C}{P^{C^2}}$$

risk ratio

$$var(\widehat{\log RR}) \approx \frac{1}{x^T} + \frac{1}{x^C}$$

odds ratio

$$\operatorname{var}(\widehat{\log OR}) \approx \frac{1}{x^{T}} + \frac{1}{P^{T} - x^{T}} + \frac{1}{x^{C}} + \frac{1}{P^{C} - x^{C}}$$

What are the problems with rare events trials?

problems can occur on two levels with zero-studies

disturbances with weighted average computation:

$$\overline{\log RR} = \frac{\sum_{i} w_{i} \widehat{\log RR_{i}}}{\sum_{i} w_{i}}$$

$$w_{i} = \frac{1}{var(\widehat{\log RR})}$$
in a similar way for *RD* and *OR*

Meta-Analysis of Rare Events

Effect estimation using Mantel-Haenszel

Contents

- What are rare events data?
- What are the problems with rare events trials?
- Effect estimation using Mantel-Haenszel
- Poisson with fixed and random effects
- Sensitivity analysis: the effect of excluding zero-studies

<ロト<日、<日、<日、<日、<日、<日、<日、<日、<日、<日、<15/68

- **Zero-inflation models**
- Logistic regression modelling
- **Conditional logistic regression modelling**

Effect estimation using Mantel-Haenszel

strategies to cope with zero-studies

pooling all studies:

$$\widehat{RR}_{crude} = \frac{(\sum_{i} x_{i}^{T})/(\sum_{i} P_{i}^{T})}{(\sum_{i} x_{i}^{C})/(\sum_{i} P_{i}^{C})}$$

<ロト<日、<日、<日、<日、<日、<日、<日、<日、<日、<日、<日、<日、<16/68

 disadvantage: potentially strong confounding effect by ignoring study factor ►

Effect estimation using Mantel-Haenszel

Mantel-Haenszel

$$\widehat{RR}_{\mathsf{MH}} = \frac{\sum_{i} x_{i}^{\mathsf{T}} P_{i}^{\mathsf{C}} / P_{i}}{\sum_{i} x_{i}^{\mathsf{C}} P_{i}^{\mathsf{T}} / P_{i}},$$

where $P_i = P_i^C + P_i^T$

advantage: estimator is not sensitive to zero-studies

is also a weighted estimator

$$\frac{\sum_{i} w_{i} \widehat{RR}_{i}}{\sum_{i} w_{i}}$$

<ロト<部ト<達ト<差ト 17/68

using the weights $w_i = x_i^C P_i^T / P_i$

Effect estimation using Mantel-Haenszel

Table: Mantel-Haenszel estimate in the rare events meta-analysis of Rosiglitazone

method	estimate	confidence interval
MI		
crude	1.2561	0.9928 - 1.5911
MH	1.2782	1.0125 – 1.6137
CV		
crude	1.1281	0.8496 - 1.4987
MH	1.0257	0.7760 – 1.3557

Effect estimation using Mantel-Haenszel

_		atistics User Window Help					
ð• 🖬 E	🕽 • 📕 🖻 • 🔟	Summaries, tables, and tests					
view		Linear models and related		.026862		0	(exact)
Comm	hand	Binary outcomes		. 0262515		0	(exact) (exact)
	reatment PersonTim	1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1	455	.2788328	6.528574	1.964286	(exact)
		Ordinal outcomes	- * L .				
	eatment PersonTime	Categorical outcomes	- F - F			0	(exact) (exact)
	eatment PersonTime	Count outcomes				š	(exact)
edit				.008248			(exact)
		Exact statistics	- F	.0118036			(exact)
		Endogenous covariates	- F - 1			0	(exact) (exact)
		Sample selection models				š	(exact)
		Multilevel mixed-effects models					
			297	.0757357	4.85803	1.673179	(exact)
		Generalized linear models	- F - F	. 0080698		× ×	(exact) (exact)
		Nonparametric analysis	• i	.0000030	29,45619	. 5697074	(exact)
			503	.2296565	9.267201	1.486527	
		Time series	• 0		38.70455	. 5019011	(exact)
		Multivariate time series	* 422	.1069507	121,2051	.4932886	(exact)
						0	(exact)
		Longitudinal/panel data	• 0		39.14029	.4991023	
riables		Survival analysis) 545	4749738	3.099505	0 5,000949	(exact)
ame	Label	Epidemiology and related		C analysis	51055505		(exact)
udv		Epidemiology and related				^	Counces
		Survey data analysis	• Ti	bles for epidemiolo	gists 🕨	Incidence-rate r	ratios
ii.	Mit			ther	•	Incidence-rate	ratios calculator
,	CVt	Multivariate analysis	•			Cohort study ris	sk ratio etc.
eatment	CH I	Power and sample size) 139 726	.8496456 .7760471	1.49868 1.35573	Cohort study ris	sk ratio etc. calculator
gsize uration		Resampling	- +	hi2(11) =	2.47 P	Case-control or	dds ratio dds ratio calculator
л		Postestimation					
ersonTime						Tabulate odds o	of failure by category
gPT		Other	- + I			Ratio of odds of	f failure for two categories
udie		1				Matched case-o	control studies
		Command				Manage of Second	control calculator

Meta-Analysis of Rare Events

Effect estimation using Mantel-Haenszel

40637,162 1,295455 by 40637,162 1,295455 by 40637,162 1,295455 by 40637,264 - 40637,264 - - 40637,353 - - 40637,353 - - 40637,353 - - 40637,353 - - 40637,353 - - 40637,353 - - 40637,353 - - 40637,353 - - 40637,353 - - 40637,353 - - 40637,353 - - 40637,354 - - 40637,353 - - 40637,353 - - 40637,353 - - 40637,353 - - 40637,353 - - 40637,452 - - 40637,453 - - 40637,453 - - 40637,453 - - 40637,453 - - 40637,453 - - 40637,453 - - 50 - - <	×	49653/145		ir - Cohort studies	and the second second	
by 44633/211 1.295455 by 46537/212 1.295455 by 46537/212 4 46537/212 4 46537/213 4 477/213 4 477/2		49653/147				
49653/234 44653/234 44653/235 44653/235 44653/235 44653/235 44653/235 44653/235 44653/235 44653/235 44653/235 44653/235 44653/235 71273/008 ADPT19 ADPT19 AVAID540 AUX10185 AVAID540		49653/162		Main if/in Weights Options		
by 44633/284 by 44633/284 44633/284 44633/380	by	49653/211	1.295455			
9/m 40883/284 9/m 40883/284 40883/284 • 40883/284 • 40883/284 • 40883/284 • 40883/284 • 40883/284 • 40883/285 • 40883/285 • 40883/285 • 40883/285 • 40883/285 • 40883/285 • 40883/285 • 40883/285 • 40883/285 • 40883/285 • 40883/285 • 40883/285 • 40883/285 • 8EL46952/334 • 68-71273/000 • 8E-71273/000 • 9 • 9 • 9 • 9 • 9 • 9 • 9 • 9 • <tr< td=""><td></td><td>49653/234</td><td></td><td>Case variable: Expos</td><td>ed variable:</td><td></td></tr<>		49653/234		Case variable: Expos	ed variable:	
40633/320 4063/330 4063/330 4063/330 4063/340 4064/36 40669 4064/36 40669 400669 4064/36 40669 40669 40669 40669 40669 400		49053/282		hu Davat		
49637,330 49637,330 40637,352 40637,359 40637,359 40637,359 40637,359 40637,359 40637,359 40737,359 40737,359 40737,359 40737,357,00 58-712737,000 58-7127,000 58-712737,000 58-712737,000 58-7127,000	, by	49033/284		pv 💌 read	nerk 💌	
40653/331 40653/332 40637/352 40637/352 40637/352 712737/008 ACOPTID - 7932397 ACOPTID - 7932397 ACOPTID - 793297 ACOPTID - 7932		49033/323		Barran Kara andaklar		
40633/350 40633/360 105		49653/330				
40633,300 1210503,330 1210503,330 1210503,330 AD00719 AD0071		49652/251		Person Time 💌		
40633/452 712737.008 A000713 A000713 A000713 A000713 A000713 A000713 A000713 A000713 A000713 A000713 A000712 A000700513 A000712 A00		49653/369				
712733,008 A007138 A007138 A007138 A007138 A007138 A007138 A007138 A007138 A007138 A007138 A007138 A00718		49653/452				
AbD/FT19 .795327 AKAU02198 .795327 AKAU02198		712753/008				
ARAL02198 AVA100193 AVA10918 AVA10018 AVA10018 AVA10018 AVA10018 AVA10018 AVA10018 AVA10018 AVA10018 AVA10018 AVA10018 AVA10018 AVA10 AVA10018 AVA10018 AVA10018 AVA10018 AVA10018 AVA1000			7953297			
AVA105610 1.35750 AVA10512 AVA10512 AVA10512 AVA10252 BRL49635/33		ARA102198				
AVD1005210 AVD107209 V70104742 BRL49655x/337 K BRL49655x/337 K BRL4965x/337 K BRL496x/37 K BRL4965x/337 K BRL4965x/337 K BRL4965		AVA100193				
AUD 02209 AVD 02209 AVD 02428 BRL 4953C / 334 BRL 4952C / 334 BRL 495		AVA105640				
AVD104742 Ref 49512 / 132 Ref 49512 / 133 Ref 495688 M-H combined 1.055728 .7760471 1.355733 (exact)		AVD100521				
Антороски 2.054422 BRL4963C/385 BRL4963C/334 BRL4963C/334 0 BRL4963C/334 0 SB-712753/002 .1.09545 SB-712753/003 SB-712753/003 SB-712753/003 SB-712753/002 Concel Sdmit M-H combined 1.025726 M-H combined 1.025726 -7760471 1.355733			0			
> BRL4953C/35 i × BRL4953C/35 i BRL4953C/37 1.195451 SE-71273/000						
BRL49652C/337 DELAM BRL49652C/337 DELAM 1.1095X51 5B-712753/000 Crude 1.128139 Crude 1.128139 M-H combifred 1.028728 .7768471 1.355733 (exact) (exact) (exact)			2.054422			
x BRL49632(337) 1.109545 1.00545 1.109545 1.00545 1.00545 1.00545 1.00545 1.00545 1.00545 0 © © OK Canal Submit 0 (exact) 1.005726 0 (exact) 0 (exact) 0 (exact) 0 (exact)	+	BRL49653C/185				
DREAM SEP-71273/000 1.199545 3477351 OK Cancel Submt SEP-71273/000 -		BRL49653C/334	0			
всовр 58-71273/002	×		4 400545			
SB-712733/002 SB-712733/001 SB-712733/009 SB-712733/009 SB-712733/009 CK Cancel Sdow 0 (Exact) C (Cancel Sdow 0 (Exact) C (1.		1.199040			
sm-712733/007 sm-712733/007 sm-712733/007 м-н combined и.1.228726 инн combined и.1.228726 инн combined и.1.228726 инн combined и.1.228726 инн сомbined и.1.228726 инн сомbined и.228726 инн сомbined и.1.228726 инн сомbined и и сомbined и и сомbined и и сомbined и и сомbined и и сомbined и и сом			.0477331	000		
SB-712733/007 SB-712733/009 Crude 1.128139 .8496456 1.498689 M-H combined 1.025726 .7760471 1.355733 (exact)	1	58-712753/003		C U H	UK	Cancel Submt
SB-712733/009 • 0 (exact) Cruds 1.178139 .8496456 1.498689 (exact) M-H combined 1.035728 .7760471 1.555733 (exact)	i.	SB-712753/007				
M-H combined 1.025726 .7760471 1.355733	1	SB-712753/009			0 (exact)	
M-H combined 1.025726 .7760471 1.355733	1.1	Crude	1,128139	. 8496456 1. 498689	(exact)	
	i i					
	1			i2(11) = 2.47 Pr>chi2 =		
		- 11-				
	111					
1. edit	1	- preserve				
- preserve	1					
		Command				
- preserve -						

Meta-Analysis of Rare Events

Effect estimation using Mantel-Haenszel

	49653/145		ir - Cohort studies		×
	49653/147 49653/162		Main if/in Weights Options		
	49653/211	1.295455	Main if/in Weights Options		
ne, by	49653/234	1.299433			
ne, by	49653/282		Stratify on variable:	Test-based conf. intervals	
ne, by	49653/284		study		
ie, by	49653/325		uody 💽	95 Confidence level	
	49653/330		Within-stratum weights		
	49653/331		and the second se		
	49653/351		Use Mantel-Haenszel		
	49653/369		C Use external		
	49653/452		O Use internal		
	712753/008				
	ADOPT19	.7953297	User-specified variable:		
	ARA102198		· ·		
	AVA100193				
	AVA105640				
	AVD100521	1.357503	Display pooled estimate		
	AVD102209		Display pooled escillate		
	AVD104742		No crude estimate		
	AVM100264	2.054422			
, BRL49653C/1			No homogeneity test		
	BRL49653C/334		Calculate std. incidence-rate difference		
×	BRL49653C/337		Calculate std. Incidence-rate difference		
11.	DRE.AM	1.199545 .8477351			
	RECORD	.8477351	~~~~		
1	SB-712753/002		00	OK Cancel Sub	nt
	SB-712753/003				
	5B-712753/007 5B-712753/009			0 (exact)	
	58-/12/53/009			(exact)	
10	Crude	1.128139	.8496456 1.498689		
11	M-H combined	1.025726	.7760471 1.355733		

►

Effect estimation using Mantel-Haenszel

Mantel-Haenszel for OR

$$\widehat{OR}_{\mathsf{MH}} = \frac{\sum_{i} x_{i}^{\mathsf{T}} (P_{i}^{\mathsf{C}} - x_{i}^{\mathsf{C}}) / P_{i}}{\sum_{i} x_{i}^{\mathsf{C}} (P_{i}^{\mathsf{T}} - x_{i}^{\mathsf{T}}) / P_{i}},$$

where
$$P_i = P_i^C + P_i^T$$

▶ is also a weighted estimator

$$\frac{\sum_{i} w_{i} \widehat{OR}_{i}}{\sum_{i} w_{i}}$$

using the weights $w_i = x_i^C (P_i^T - x_i^T) / P_i$

►

Effect estimation using Mantel-Haenszel

Mantel-Haenszel for RD

$$\widehat{RD}_{\mathsf{MH}} = \frac{\sum_{i} (x_{i}^{\mathsf{T}} P_{i}^{\mathsf{C}} - x_{i}^{\mathsf{C}} P_{i}^{\mathsf{T}}) / P_{i}}{\sum_{i} (P_{i}^{\mathsf{T}} P_{i}^{\mathsf{C}} / P_{i})},$$

where
$$P_i = P_i^C + P_i^T$$

▶ is also a weighted estimator

$$\frac{\sum_{i} w_{i} \widehat{RD}_{i}}{\sum_{i} w_{i}}$$

using the weights $w_i = (P_i^T P_i^C)/P_i$

-Effect estimation using Mantel-Haenszel

testing homogeneity of effect

major difficulties with Mantel-Haenszel lies in establishing homogeneity of effect

$$\chi_{k-1}^{2} = \sum_{i} \frac{\left(\widehat{\log RR}_{i} - \log \widehat{RR}_{MH}\right)^{2}}{var(\widehat{\log RR}_{i})}$$

where k is the number of studies

- this statistic will not work in the case of zero-studies
- this question needs to be approached in a modelling framework

Meta-Analysis of Rare Events

- Poisson with fixed and random effects

Contents

What are rare events data?

What are the problems with rare events trials?

Effect estimation using Mantel-Haenszel

Poisson with fixed and random effects

Sensitivity analysis: the effect of excluding zero-studies

Zero-inflation models

Logistic regression modelling

Conditional logistic regression modelling

-Poisson with fixed and random effects

Poisson regression

consider number of events X as a Poisson count with mean

 $E(X) = \mu P$

- clearly, $\mu = E(X)/P$ is the **incidence risk**
- write in study i

$$E(X_{ij}) = \mu_j P_{ij}$$

<ロト<回ト<三ト<三ト<三ト 26/68

for j = 1 (treatment) and j = 0 (control)

• so that again $RR = \mu_1/\mu_0$

- Poisson with fixed and random effects

Poisson regression

▶ in study *i*

$$E(X_{ij}) = \mu_j P_{ij}$$

take logarithms on both sides

$$\log E(X_{ij}) = \log P_{ij} + \log \mu_j = \log P_{ij} + \alpha + \beta \times j$$

- so that $\beta = \log(\mu_1/\mu_0)$ is the log-risk ratio
- log P_{ij} enters as a covariate with known coefficient into the model: an offset

Poisson with fixed and random effects

	· 🖻 🖸 💿 🔇	Thep	
×	Istudie 38	2.08e-16 1.86e-08	-0.00 1.000 0 .
	_Istudie_39	.9590241 1.356372	-0.03 0.976 .0599727 15.33577
	_Istudie_40	1.026627 1.451927	poisson - Poisson regression
onTime, by	_Istudie_41 _Istudie_42	.6262025 .8857064 1.71e-16 1.51e-08	E poisson - Poisson regression
nTime, by	_Istudie_42	4.808375 5.14171	Model by/if/in Weights SE/Robust Reporting Max options
nTime, by	_Istudie_44	.84894 1.200641	Podda byzzni wegrus szchobás hepoting más opions
innine, by	Istudie 45	6.85e-16 6.12e-08	Dependent variable: Independent variables:
	_Istudie_46	4.11e-16 3.65e-08	
, exposure	_Istudie_47	1.81e-16 1.60e-08	_ cv v treatment
ment, exp	_Istudie_48	8.16e-16 7.22e-08	Suppress constant term
ment, exp	_Istudie_49	1.72e-16 1.53e-08 4.66e-16 4.15e-08	
ent i.study	_Istudie_50 _Istudie_51	4.66e-16 4.15e-08 5.45e-16 4.86e-08	Options
en istudy	_Istudie_52	1.26e-15 1.13e-07	
	_Istudie_53	1.34e-15 1.20e-07	Exposure variable: Offset variable:
ent i.studi	_Istudie_54	.3438387 .3516613	Person Time 👻 🗸
ent i.studi	_Istudie_55	.0997788 .1067717	
t, exposure	_Istudie_56	1.455544 1.461502	Constraints:
	PersonTime	(exposure)	· · · · · · · · · · · · · · · · · · ·
,	. poisson cv t	reatment, exposure(Pers	onTrir Keep collinear variables (rarely used)
×	Iteration 0:	log likelihood172.	0608
	Iteration 1:	log likelihood = -172.	0216
<u></u>	Iteration 2:	log likelihood = -172.	0216
E	Poisson regres		
			OK Cancel Submit
	Log likelihood	= -172,02165	
	cv	IRR Std. Err.	z P> z [95% Conf. Interval]
	treatment	1.128139 .1576322	0.86 0.388 .8578792 1.483539
	PersonTime	(exposure)	
=2			
-3	Command		
=4 -			

<ロト

- Poisson with fixed and random effects

Poisson regression with random study effect

taking into account the study effect:

the effect homogeneity model

$$\log E(X_{ij}) = \log P_{ij} + \alpha_i + \beta \times j$$

the effect heterogeneity model

$$\log E(X_{ij}) = \log P_{ij} + \alpha_i + \beta_i \times j$$

- Poisson with fixed and random effects

Poisson regression with random study effect

two options:

- fixed effects model: α_i and β_i are treated as fixed parameters
- ▶ disadvantage: many studies → many parameters
- Neyman-Scott problem (sample size and number of parameters connected)
- random effects model: α_i and β_i are treated as random quantities:

$$\alpha_i \sim N(\alpha, \sigma_{\alpha}^2)$$
 and $\beta_i \sim N(\beta, \sigma_{\beta}^2)$

Poisson regression with random study effect

this leads to the following different likelihoods (in the example of the ${\color{black} \textbf{homogeneity model}}$)

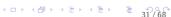
► fixed effects model:

$$\prod_{i} \left[Po(x_{i0}|P_{i0} \exp(\alpha_i)) \times Po(x_{i1}|P_{i1} \exp(\alpha_i + \beta)) \right], \quad (1)$$

random effects model:

$$\prod_{i} \int [Po(x_{i0}|P_{i0}\exp(\alpha_i))]$$

 $\times Po(x_{i1}|P_{i1}\exp(\alpha_i+\beta))]\phi(\alpha_i|\alpha,\sigma_{\alpha}^2)d\alpha_i.$



- Poisson with fixed and random effects

Poisson regression with random study effect

likelihood in the example of the heterogeneity model

► random effects model:

$$\prod_{i} \int Po(x_{i0}|P_{i0}\exp(\alpha_i)) \times$$

$$\left[\int \mathsf{Po}(\mathsf{x}_{i1}|\mathsf{P}_{i1}\exp(\alpha_i+\beta_i))\phi(\beta_i|\mathsf{0},\sigma_\beta^2)\mathsf{d}\beta_i\right]\phi(\alpha_i|\alpha,\sigma_\alpha^2)\mathsf{d}\alpha_i.$$

- Poisson with fixed and random effects

Poisson regression with random study effect

integrals have no closed form solution:

- Laplace approximation
- Gauss-Hermite quadrature

Poisson with fixed and random effects

onTime, by	Refining start		Î	stmepoisson Multilevel mixed-effects Poisson regression							
onTime, by onTime, by onTime, by t, exposure	Iteration 0: log likelihood = -102.75699 Iteration 1: log likelihood = -100.42017 Iteration 2: log likelihood = -100.38951 Performing gradient-based optimization:			Model Integration by/K/in Reporting Max options Dependent variables: independent variables: independent variables: independent variables:							
tment, exp tment, exp ient i.study	Iteration 0: Iteration 1: Iteration 2: Iteration 3:	log likelihood log likelihood log likelihood log likelihood	<pre>-100.30978 -100.30949 -100.30949</pre>	Exposure / 0	fset						
ent i.studi ent i.studi t, exposure	Mixed-effects Group variable	Poisson regress : study	lon	Exposure Person Time Random-effer		[) Offset vari	able:			
tment, exp	Integration po Log likelihood	ints = 7		Level equation	Level variable		Factor equation	Factor variable/ Independent variables		Con stru	
×				EQ 1	study	•				inc	
	CV		td. Err. :	EQ 3					*	inc	
E	treatment PersonTime	1.019212 .: (exposure)	1432954 0.3	EQ 4						in	
				EQ 5		-			w.	1	
	Random-effec	ts Parameters	Estimate	EQ 6		*			Ŧ	in	
	study: Identit	y var (_cons)	1.229417	EQ 7		Y			¥.	ind	
	LR test vs. Po	isson regressio	n: chibar2(01)	00 1						_	

Testing homogeneity with the likelihood ratio test

► random effects model *M*₁:

$$L_1 = \prod_i \int Po(x_{i0}|P_{i0}\exp(\alpha_i))Po(x_{i1}|P_{i1}\exp(\alpha_i+\beta))\phi(\alpha_i|0,\sigma_{\alpha}^2)d\alpha_i$$

<ロ > < 回 > < 国 > < 国 > < 国 > < 国 > < 国 > 35/68

► NO random effects *M*₀:

$$L_0 = \prod_i \left[Po(x_{i0} | P_{i0} \exp(\alpha)) \times Po(x_{i1} | P_{i1} \exp(\alpha + \beta)) \right]$$

likelihood ratio

$$\log \lambda = 2 \log L_1 / L_0$$

is χ^2 with 1 df under the M_0

-Poisson with fixed and random effects

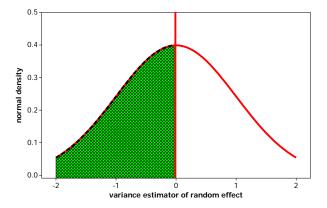
Testing homogeneity with the likelihood ratio test

variance estimates cannot be negative :

$$\hat{\sigma}_{\alpha}^2 \geq 0$$

• hence: distribution of $\hat{\sigma}_{\alpha}^2$ cannot be normal

-Poisson with fixed and random effects



Testing homogeneity with the likelihood ratio test

asymptotic distribution:

$$P(\hat{\sigma}_{\alpha}^2/s.e.(\hat{\sigma}_{\alpha}^2) < x) = 0.5 + 0.5\Phi(x)$$

where $\Phi(x)$ is the CDF of a standard normal distribution

similarly for the asymptotic distribution of the likelihood ratio

$$\log \lambda = 2 \log L_1 / L_0 \sim 0.5 + 0.5 \chi^2_{(1)}$$

in practice, conventionally computed P-values need only be divided by 2 since:

 $P(\log \lambda > \log \lambda_{obs}) = 1 - [0.5 + 0.5\{1 - P(\log \lambda > \log \lambda_{obs,old})\}]$ $= 0.5P(\log \lambda > \log \lambda_{obs,old})$

Poisson with fixed and random effects

Table: Poisson regression estimates in the rare events meta-analysis of Rosiglitazone; Log-L stands for the maximised log-likelihood

Poisson model	estimate	confidence interval	Log-L		
MI					
treatment	1.2561	0.9991 - 1.5793	-174.2054		
treatment	1.2634	1.0006 - 1.5952	-137.9558		
σ_{lpha}^2	0.6346				
treatment	1.2634	1.0006 - 1.5952	-137.9558		
σ_{lpha}^2	0.6346				
σ_{β}^2	0.				

Poisson with fixed and random effects

Table: Poisson regression estimates in the rare events meta-analysis of Rosiglitazone; Log-L stands for the maximised log-likelihood

Poisson model	estimate	confidence interval	Log-L	
CV				
treatment	1.1281	0.8579 – 1.4835	-172.0216	
treatment	1.0192	0.7737 - 1.3426	-100.3095	
σ_{lpha}^2	1.2294			
treatment	1.0192	0.7737 - 1.3426	-100.3095	
σ_{α}^2	1.12294			
σ_{β}^2	0.			

model evaluation

- for model assessment we will use criteria that compromise between model fit and model complexity
- Akaike information criterion

$$AIC = -2\log L + 2p$$

Bayesian Information criterion

$$BIC = -2\log L + p\log k$$

- where p is the number of parameters in the model
- ▶ and *k* is the number of trials in the meta-analysis
- we seek a model for which AIC and/or BIC are small

-Sensitivity analysis: the effect of excluding zero-studies

Contents

- What are rare events data?
- What are the problems with rare events trials?
- **Effect estimation using Mantel-Haenszel**
- Poisson with fixed and random effects
- Sensitivity analysis: the effect of excluding zero-studies

- Zero-inflation models
- Logistic regression modelling
- Conditional logistic regression modelling

Sensitivity analysis: the effect of excluding zero-studies

sensitivity analysis:

how does the effect estimate of the risk ratio depend on the exclusion/inclusion of

- double-zero (DZ)
- single-zero (SZ)

studies?

-Sensitivity analysis: the effect of excluding zero-studies

Table: Poisson random effects regression estimates of the risk ratio: the effect of excluding double-zero (DZ) and single-zero (SZ) studies and none excluded (NONE); number of studies included is given in brackets in the first column

excluding (k)	RR	SE	Ζ	P-value	95% CI
MI					
NONE(56)	1.2633	0.1503	1.96	0.049	1.0006 - 1.5952
DZ(41)	1.2634	0.1503	1.97	0.049	1.0008 - 1.5955
SZ(15)	1.2101	0.1512	1.53	0.127	0.9473 – 1.5458
CV					
NONE(56)	1.0193	0.1433	0.14	0.892	0.7738 - 1.3426
DZ(27)	1.0246	0.1441	0.17	0.863	0.7778 – 1.3497
SZ(8)	0.9427	0.1395	-0.40	0.690	0.7054 - 1.2599

Contents

What are rare events data?

What are the problems with rare events trials?

Effect estimation using Mantel-Haenszel

Poisson with fixed and random effects

Sensitivity analysis: the effect of excluding zero-studies

<ロ > < 団 > < 臣 > < 臣 > 三 900 45/68

Zero-inflation models

Logistic regression modelling

Conditional logistic regression modelling

-Zero-inflation models

Zero-inflation models

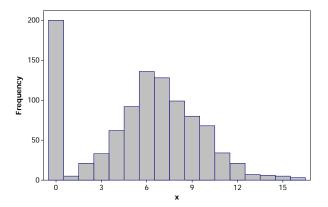
count data with many zeros lead to the question:

- ▶ is there an excess of zero counts relative to the Poisson model
- an excess in zero-counts is called zero-inflation

$$Pr[X = 0] = \pi + (1 - \pi)Po(0|\mu)$$
 (2)

$$Pr[X = x] = (1 - \pi)Po(x|\mu) \text{ for } x = 1, 2, ...$$
 (3)

<ロト<部ト<差ト<差ト<差ト 46/68 Zero-inflation models



-Zero-inflation models

Zero-inflation models

Lambert (1992) extended the simple ZIP-model to covariates:

$$\log \mu_{ij} = \log P_{ij} + \alpha + \beta \times j \tag{4}$$

logit
$$\pi_{ij} = \log \frac{\pi_{ij}}{1 - \pi_{ij}} = \alpha' + \beta' \times j.$$
 (5)

Zero-inflation models

		(inclu			
• 🔟 • 🔟	- 🗖 🗖 💿 😒				
×	Iteration 8: Iteration 9: Iteration 10:	log likelihood = -172.39408 log likelihood = -172.39408 log likelihood = -172.39408 log likelihood = -172.39408	(not concave) (not concave) (not concave) (not concave)		
nent i.studi	Iteration 11: Iteration 12:	log likelihood = -172.39408 log likelihood = -172.39408	zip - Zero-inflated Pois	sson regression	
nent i.studi	Iteration 13:	log likelihood = -172,39408	Model Invition Meinhan	SE/Robust Reporting Max options	
nt, exposure.	Iteration 14: Iteration 15:				
itment, exp	Iteration 16:	log likelihood = -172.39022	Dependent variable:	Independent variables:	
late(treatm.	Iteration 17: Iteration 18:	log likelihood = -172.37993	cv 💌	treatment	6
late(treatm.	Iteration 19: Iteration 20:			Suppress constant term	
late(_cons).			Inflation variables:	Constant inflation	
late(treatm.	Fitting full m	model:	treatment		
late(treatm.	Iteration 0:	log likelihood = -172.37984	Options		
late(treatm.	Iteration 1: Iteration 2:	log likelihood = -171.9954 log likelihood = -171.99514	Exposure variable:	Offset variable:	
late(_cons).	Iteration 3:	log likelihood = -171.99514	Person Time		
late(_cons). +	zero-inflated	Poisson regression			
×			Inflation offset variable:	Constraints:	
^				V	-
ñ	Inflation mode	el = logit d = -171.9951	Keep collinear variab	oles (rarely used)	
E			Model for characterizin	ng zeros	
_		IRR Std. Err.	z 🔍 logt	probit	
	cv treatment	1.131013 .1588599 0.	88		
	PersonTime	(exposure)	00	OK Cancel	Submit
	inflate				
	_cons	-3.522325 4.462084 -0.	79 0.430 -12.2678	5 5.223199	
	•				
:=2					
:=3	Command				
=4 +					
+					

Contents

- What are rare events data?
- What are the problems with rare events trials?
- **Effect estimation using Mantel-Haenszel**
- Poisson with fixed and random effects
- Sensitivity analysis: the effect of excluding zero-studies

<ロト<部ト<差ト<差ト 50/68

- Zero-inflation models
- Logistic regression modelling
- Conditional logistic regression modelling

Logistic regression

let Y_{ij} denote the binary outcome for an event ($Y_{ij} = 1$) in study i and treatment arm j (j = 0, 1)

$\pi_{ij} = P(Y_{ij} = 1)$ probability of an event

logistic transformation

$$\log \frac{\pi_{ij}}{1 - \pi_{ij}} = \alpha + \beta \times j$$

so that β is the log-odds ratio

Logistic regression model

 each trial arm within each study contributes a binomial likelihood

$$\binom{n_{ij}}{x_{ij}}\pi^{x_{ij}}_{ij}(1-\pi_{ij})^{n_{ij}-x_{ij}}$$



$$\log \frac{\pi_{ij}}{1 - \pi_{ij}} = \alpha + \beta \times j$$

□ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ >

Logistic likelihood

 $L = \prod_{i} \prod_{j} {n_{ij} \choose x_{ij}} \pi_{ij}^{x_{ij}} (1 - \pi_{ij})^{n_{ij} - x_{ij}}$

where

$$\log \frac{\pi_{ij}}{1 - \pi_{ij}} = \alpha + \beta \times j$$

Logistic regression with random intercept effect for study

 $\log \frac{\pi_{ij}}{1 - \pi_{ij}} = \alpha_i + \beta \times j$

• where $\alpha_i \sim N(\alpha, \sigma_{\alpha}^2)$ is a random intercept effect

►

Logistic regression modelling

Mixed Logistic Likelihood

$$L = \prod_{i} \int_{\alpha_{i}} \prod_{j} {n_{ij} \choose x_{ij}} \pi_{ij}^{x_{ij}} (1 - \pi_{ij})^{n_{ij} - x_{ij}} \phi(\alpha_{i}) d\alpha_{i}$$

where φ(α_i) is a normal density with mean α and variance σ²_α
 and π::

$$\log \frac{\pi_{ij}}{1 - \pi_{ij}} = \alpha_i + \beta \times j$$

Logistic regression with random intercept effect for study

 $\log \frac{\pi_{ij}}{1 - \pi_{ij}} = \alpha_i + \beta_i \times j$

• where $\alpha_i \sim N(\alpha, \sigma_{\alpha}^2)$ is a random intercept effect

▶ and $\beta_i \sim N(\beta, \sigma_{\beta}^2)$ is a random slope (treatment) effect

Mixed Logistic Likelihood

$$L = \prod_{i} \int_{\alpha_{i}} \left(\int_{\beta_{i}} \prod_{j} \binom{n_{ij}}{x_{ij}} \pi_{ij}^{\mathbf{x}_{ij}} (1 - \pi_{ij})^{n_{ij} - \mathbf{x}_{ij}} \phi(\beta_{i}) d\beta_{i} \right) \phi(\alpha_{i}) d\alpha_{i}$$

where φ(α_i) is a normal density with mean α and variance σ²_α
 where φ(β_i) is a normal density with mean β and variance σ²_β
 and π::

$$\log \frac{\pi_{ij}}{1 - \pi_{ij}} = \alpha_i + \beta_i \times j$$

Mixed-effects Binomial varia				Number of	obs =	36
Group variable		n ly_02		Number of	groups =	18
				Obs per g	roup: min = avg = max =	2.0
Integration me	ethod: mvaghe	rmite		Integrati	on points =	7
Log likelihood	d = -78.215149	9		Wald chi2 Prob > ch		
x	Odds Ratio	Std. Err.	z	P> z	[95% Conf.	Interval]
Treat_bin	.2458528	.0857953	-4.02	0.000	.1240598	.4872135
_cons	.0310596	.0077302	-13.95	0.000	.0190698	.0505877
study 02						
var(Treat ~n)	.6187426	.5829695			.0976161	3.921919
var(_cons)	.7717062	.3642633			.3059603	1.946431
LR test vs. lo	gistic regres	ssion: d	chi2(2) =	50.11	Prob > chi	2 = 0.0000

Table: Logistic regression estimates in the rare evidence meta-analysis of CRBSI; Log-L stands for the maximised log-likelihood

logistic model	estimate	confidence interval	Log-L
treatment	0.30	0.20 - 0.47	-103.27
treatment	0.29	0.19 - 0.46	-79.70
σ_{lpha}^2	0.74	0.30 - 1.87	
treatment	0.25	0.12 - 0.49	-78.22
σ_{α}^2	0.77	0.31 - 1.95	
σ_{β}^2	0.62	0.10 - 3.92	

- Conditional logistic regression modelling

Contents

- What are rare events data?
- What are the problems with rare events trials?
- Effect estimation using Mantel-Haenszel
- Poisson with fixed and random effects
- Sensitivity analysis: the effect of excluding zero-studies
- Zero-inflation models
- Logistic regression modelling

Conditional logistic regression modelling

Conditional logistic regression modelling

recall:

let
$$RR = \mu_1/\mu_0$$
 and $X_i = X_{i1} + X_{i0}$

▶ in study *i*, for treatment

$$E(X_{i1}) = \mu_1 P_{i1}$$

for control

$$E(X_{i0}) = \mu_0 P_{i0}$$

Conditional logistic regression modelling

it follows:

• then $E(X_{i1} + X_{i0}) = \mu_1 P_{i1} + \mu_0 P_{i0}$ so that

$$E(X_{i1}|X_i) = X_i \frac{\mu_1 P_{i1}}{\mu_1 P_{i1} + \mu_0 P_{i0}} = X_i \frac{RR \frac{P_{i1}}{P_{i0}}}{1 + RR \frac{P_{i1}}{P_{i0}}}$$

depends only on RR, the parameter of interest

Conditional logistic regression modelling

Table: Layout for conditional logistic regression in study *i*

	treatment contro		margin
events	X_{i1}	X_{i0}	Xi
person time	P_{i1}	P_{i0}	Pi

$$X_{i1}|X_i \sim Bin(q_i, X_i)$$
 with $q_i = \frac{\mu_1 P_{i1}}{\mu_1 P_{i1} + \mu_0 P_{i0}} = \frac{RR\frac{P_{i1}}{P_{i0}}}{1 + RR\frac{P_{i1}}{P_{i0}}}$

Conditional logistic regression modelling

furthermore:

► let
$$RR = \exp(\beta)$$

$$q_i = \frac{RR\frac{P_{i1}}{P_{i0}}}{1 + RR\frac{P_{i1}}{P_{i0}}}$$

$$= \frac{\exp[\beta + \log(\frac{P_{i1}}{P_{i0}})]}{1 + \exp[\beta + \log(\frac{P_{i1}}{P_{i0}})]}$$

$$\frac{q_i}{1 - q_i} = \exp[\beta + \log(\frac{P_{i1}}{P_{i0}})]$$

$$\log\left(\frac{q_i}{1 - q_i}\right) = \beta + \log(\frac{P_{i1}}{P_{i0}})$$

- Conditional logistic regression modelling

hence:

►

$\log\left(rac{q_i}{1-q_i} ight)=eta+\log(rac{P_{i1}}{P_{i0}})$

▶ we find RR as logistic regression with intercept only and offset log(P_{i1}/P_{i0})

• note that β is a log-risk ratio

- Conditional logistic regression modelling

Table: Meta-analysis on rare evidence data on the effect of anti-infective-treated catheter in compariosn to standard catheter; CRBSI (X_{i1}, X_{i0}) refers to catheter-related bloodstream infection events, n_{i1} , n_{i0} is the size of the respective study arm

	control arm		treatment arm		conditional	
study ID	X _{i0}	n _{i0}	<i>X</i> _{<i>i</i>1}	n _{i1}	Xi	n _{i1} /n _{i0}
1	3	117	0	116	3	116/117
2	3	35	1	44	4	44/35
3	9	195	2	208	11	208/195
4	7	136	0	130	7	130/136
5	6	157	5	151	11	151/157
6	4	139	1	98	5	98/139
7	3	177	1	174	4	174/177
8	2	39	1	74	3	74/39

Conditional logistic regression modelling

```
. melogit xt, offset(log ratio) binomial(xsum) or
Iteration 0:
             log likelihood = -26.454133
Iteration 1: log likelihood = -26.199518
Iteration 2: log likelihood = -26.199183
Iteration 3:
               log likelihood = -26.199183
Logistic regression
                                                 Number of obs
                                                                             18
                                                                     =
Binomial variable:
                           ¥ S11m
                                                 Wald chi2(0)
                                                                     =
Log likelihood = -26.199183
                                                 Prob > chi2
                                                                    =
               Odds Ratio
                                                           [95% Conf. Interval]
          xt.
                            Std. Err.
                                                 P>|z|
                                            z
       cons
                 .3072359
                            .0678268
                                         -5.35
                                                 0.000
                                                           .1993228
                                                                       .4735732
   log ratio
                        1
                           (offset)
```

Conditional logistic regression modelling

can be easily extend to random effects model

$$\log\left(rac{q_i}{1-q_i}
ight)=eta_i+\log(rac{P_{i1}}{P_{i0}})$$

with $\beta_i \sim N(\beta, \sigma_{\beta}^2)$

Mixed-effects Binomial vari		Number o	=	18			
Binomial variable: xsum Group variable: study_short				Number o	of groups	=	18
				Obs per	group: min	=	1
					avg	=	1.0
					max	=	1
Integration m	ethod: mvaghe	rmite		Integrat	tion points	=	7
Log likelihoo	d = -24.79442	9		Wald chi Prob > c		= =	
xt	Odds Ratio	Std. Err.	z	₽> z	[95% Con	f.	Interval]
_cons log_ratio	.2718073 1	.0920478 (offset)	-3.85	0.000	.1399591		. 527863
study_short var(_cons)	. 6007426	.5915841			.0871894		4.139168

LR test vs. logistic regression: chibar2(01) = 2.81 Prob>=chibar2 = 0.0469