

階層ベイズモデルとその応用

野間 久史
統計数理研究所
2023年12月14日
日本計量生物学会 2023年度計量生物セミナー
e-mail: noma@ism.ac.jp
URL: <http://www.ism.ac.jp/~noma/>

ベイズ統計学 Bayesian Statistics

- ▶ 統計学の学派のひとつ
- ▶ 現在のスタンダードである、信頼区間やP値を用いた統計解析は、頻度論 (frequentist) という学派による方法
- ▶ 頻度論では、データから得られた情報のみを用いて、治療法の有効性などについての分析が行われる
- ▶ ベイズ統計学では、事前情報を事前分布 (prior distribution) として解析に組み込み、研究者の主観に基づく「主観確率」による解析を行う

2

JAMA Network
JAMA* Journals Enter Search Term

This article was corrected | View Correction

This Issue View Citations Altmetric 85

Download PDF More CME & MOC Cite This Permissions

Original Investigation
October 24/31, 2017

Effect of Therapeutic Hypothermia Initiated After 6 Hours of Age on Death or Disability Among Newborns With Hypoxic-Ischemic Encephalopathy: A Randomized Clinical Trial

Abbot R, Laptook M¹, Seetha Shankaran, MD¹, Jon E. Tyson, MD, MPH², et al.
JAMA. 2017;318(16):1550-1560. doi:10.1001/jama.2017.14972

Related Articles

Key Points
Question: Does initiation of hypothermia at 6 to 24 hours after birth reduce the risk of death or disability at 18 months among term newborns with hypoxic-ischemic encephalopathy?

JAMA Guide to Statistics and Methodsでの
ベイズ統計学の解説の題材となった研究
です。

JAMA 2017; 318(16): 1550-60.

3

研究の概要

- ▶ 妊娠期間36週以降の低酸素性虚血性脳症の新生児に、出生後6時間以内で低体温療法を開始すると、死亡・障害のリスクを減じることができることがわかってきた
- ▶ しかし、出生後6時間以後の有効性を評価した試験はなかった
- ▶ 出生後6-24時間での治療開始によって、18カ月時点での死亡・障害のリスクが減じられるかを評価したかった
- ▶ 米国21施設において、多施設共同ランダム化比較試験が行われた
 - ▶ 96時間の33.5°C (許容範囲: 33-34°C) の体温維持 (n=83)
 - ▶ 同じく 37.0°C (許容範囲: 36.5-37.3°C) の体温維持 (n=85)

Laptook et al. (2017) 4

ベイズ統計学の活用

- ▶ 試験に組み入れることができる新生児の数が限られていることから、この試験は計画段階から十分な検出力での評価が難しいという想定があった
- ▶ ベイズ流の統計解析を行うことに！！
- ▶ ベイズ推測の情報量 = データの情報量 + 事前情報
- ▶ 従来の頻度論的な統計解析で、有意差が得られないような場合でも、治療効果の評価において、臨床的に有用な結果が得られることも

Quintana et al. (2017) 5

Table 3. Primary and Secondary Outcomes: aRRs and Posterior Probability of Treatment Effect^a

Outcome	No. (%) Hypothermia (n = 78)	No. (%) Noncooled (n = 79)	Enthusiastic Prior (RR, 0.72)		Neutral Prior (RR, 1.10)		Skeptical Prior (RR, 1.10)	
			aRR (95% Credible Interval)	P-TB, %	aRR (95% Credible Interval)	P-TB, %	aRR (95% Credible Interval)	P-TB, %
Primary Outcome								
Death or moderate-severe disability	19 (24.4)	22 (27.9)	0.78 (0.52-1.15)	90	0.86 (0.58-1.29)	76	0.89 (0.60-1.32)	73
Secondary Outcomes								
Death ^b	9 (11.5)	9 (11.4)	0.74 (0.45-1.21)	89	0.86 (0.54-1.44)	73	0.90 (0.56-1.52)	67
Moderate or severe disability ^c	10 (12.8)	13 (16.5)	0.74 (0.44-1.24)	87	0.89 (0.54-1.48)	68	0.93 (0.56-1.55)	61
Severe disability ^c	9 (11.5)	12 (15.2)	0.73 (0.43-1.23)	88	0.88 (0.53-1.50)	68	0.93 (0.55-1.55)	61
Moderate disability ^c	1 (1.3)	1 (1.3)						
Mild disability ^c	16 (20.5)	12 (15.2)	1.0 (0.62-1.62)	50	1.18 (0.73-1.91)	25	1.23 (0.76-2.0)	20

Abbreviations: aRR, adjusted risk ratio; P-TB, posterior probability of treatment benefit (risk ratio < 1.0); RR, risk ratio.

^a Reference for the aRR is the noncooled group and the aRR is adjusted for level of encephalopathy (moderate, severe) and age at randomization (=12 h, >12 h). Variables in the adjusted analyses were outcome, treatment, level of encephalopathy at randomization, and age at randomization.

^b Causes of death in the hypothermia group were asphyxia brain injury (n = 5), multiorgan failure (n = 1), persistent pulmonary hypertension (n = 1), respiratory failure associated with intractable seizures (n = 1), and intracranial hemorrhage (n = 1). Causes of death in the noncooled group were asphyxia brain injury (n = 5), multiorgan failure (n = 1), and meconium aspiration syndrome (n = 1), and 2 were without an assigned cause (after discharge).

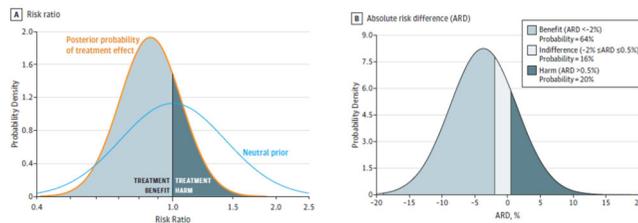
^c Disability categories were defined as follows: severe included any of the following: a cognitive score less than 70, a Gross Motor Function Classification Score (GMFCS) level of 3 to 5, or blindness or hearing impairment with inability to follow commands despite amplification; moderate included a cognitive score between 70 and 84 and any of the following: a GMFCS level of 2, an active seizure disorder (antiepileptic drugs in use), or a hearing deficit with the ability to follow commands after amplification; and mild included a cognitive score between 85 and 94 or a cognitive score of 85 or greater with any of the following: a GMFCS level of 1 or 2, seizure disorder (without medication), or a hearing deficit with the ability to follow commands without amplification.

^d Analyses could not be performed with 1 infant in each group.

Laptook et al. (2017) 6

事後分布とベイズ推測

Figure 2. Posterior Probability of Death or Disability With Hypothermia Initiated at 6 to 24 Hours After Birth vs Noncooling



Laptook et al. (2017) 7

本日のお話

- ▶ 階層モデル Hierarchical Model
 - ▶ 複数のレベルによって構成された、階層型の構造を持つ統計モデル
 - ▶ 頻度論における一般化線形混合効果モデルのようなマルチレベルモデルと同様、データ解析の単位（個人、動物など）のクラスター効果を調整した解析が可能
 - ▶ ベイズ流の解析では、複雑な構造を持つ多段階階層モデルでもMCMCを使って統一的な解析が可能であるという利点があり、特に、複雑なモデリングを要する解析で広く用いられている

8

OXFORD JOURNALS ONLINE
International Journal of Epidemiology

Articles

Being sceptical about meta-analyses: a Bayesian perspective on magnesium trials in myocardial infarction

Julian PT Higgins, David J Spiegelhalter

International Journal of Epidemiology, Volume 31, Issue 1, February 2002, Pages 96-104, <https://doi.org/10.1093/ije/d1.31.1.96>

Published: 01 February 2002 Article history

PDF Split View Cite Permissions Share

Abstract
Background There has been extensive discussion of the apparent conflict between meta-analysis and a mega-trial investigating the benefits of intravenous magnesium following myocardial infarction, in which the early trial results have been said to be 'too good to be true'.
Methods We apply Bayesian methods of meta-analysis to the trials available before and after the publication of the ISIS-4 results. We show how scepticism can be formally incorporated into an analysis as a Bayesian prior distribution, and how Bayesian meta-analysis models allow appropriate exploration of hypotheses that the treatment effect depends on the size of the trial or the risk in the control group.

<https://academic.oup.com/ije/article/31/1/96/655931>

9

心筋梗塞後のマグネシウム静注治療

Table 1 Brief summary of landmark publications from trials and meta-analyses in magnesium for acute myocardial infarction

Year	Publication	Finding concerning magnesium treatment	Conclusion
1981	First trial report ¹²	Trend toward smaller infarct size (43, later 76, patients)	'Magnesium ... is promising and deserves further study' ¹³
1991	Meta-analysis by Teo <i>et al</i> ¹⁰	55% reduction in odds of mortality (8 trials)	'further large scale trials to confirm (or refute) the findings are desirable'
1992	LIMIT-2 large trial ¹⁴	24% reduction in mortality	'a simple, safe and widely applicable treatment'
1993	Circulation editorial ¹⁵	'An effective, safe, simple and inexpensive treatment'	Recommends further trials to obtain 'a more precise estimate of the mortality benefit'
1995	ISIS-4 mega-trial ¹⁶	Non-significant adverse mortality (58 050 patients)	'Overall, there does not now seem to be any good clinical trial evidence for the routine use of magnesium'

Higgins and Spiegelhalter (2002)

心筋梗塞後のマグネシウム静脈注射が、発症後の死亡率を低減するというランダム化比較試験のエビデンスが繰り返し報告され、BMJのメタアナリシスでも、有意な効果を示された。しかし、その後の大規模ランダム化比較試験 ISIS-4試験では、まったく真逆の結果が得られ、一転して、「有効でない」という結論が導かれた有名な事例。

10

Table 2 Summary data from 15 randomized trials of intravenous magnesium for acute myocardial infarction (data from Sterne *et al*²⁸) with classical meta-analyses using the Peto method²⁹ and DeSimonian and Laird (D-L) method³⁰ based on sample log(odds ratio) and a moment estimate of between-trial variance. A test of homogeneity of odds ratios (I^2) is as described by Wood³¹

Trial	Magnesium group		Control group	
	Deaths n_1^*	Patients n_2^*	Deaths n_1^*	Patients n_2^*
Morton	1	40	2	50
Rasmussen	9	135	23	135
Smith	2	200	7	200
Alraham	1	48	1	46
Feldbach	19	150	8	148
Shechter 1990	1	59	9	56
Czerwinski	1	25	3	23
LIMIT-2	119	1190	118	1157
Fixed effect (Peto) meta-analysis of above eight trials: OR = 0.65 (95% CI: 0.51, 0.82); $I^2_{het} = 0.19$				
Random effects (D-L) meta-analysis of above eight trials: OR = 0.55 (95% CI: 0.34, 0.89)				
Berkefeld	0	22	1	21
Singh	6	76	11	75
Beretta	1	27	7	27
Goff	5	23	13	33
Thorgersen	4	130	8	122
Shechter 1995	1	107	17	108
Fixed effect (Peto) meta-analysis of above 14 trials: OR = 0.57 (95% CI: 0.46, 0.71); $I^2_{het} = 0.19$				
Random effects (D-L) meta-analysis of above 14 trials: OR = 0.47 (95% CI: 0.33, 0.68)				
ISIS-4	2216	29 011	2103	28 050
Fixed effect (Peto) meta-analysis of above 15 trials: OR = 1.01 (95% CI: 0.95, 1.07); $I^2_{het} = 0.0002$				
Random effects (D-L) meta-analysis of above 15 trials: OR = 0.53 (95% CI: 0.36, 0.77)				

Higgins and Spiegelhalter (2002)

1991年、BMJに報告された8つのRCTによるメタアナリシス (Teo *et al.*, 1991)

1995年、ISIS-4試験の結果が報告される直前までに公表されていた6つのRCTを追加したメタアナリシス

11

“良すぎた” データと懐疑的事前分布

- ▶ Yusuf and Flather (1995) の懐疑的意見「このケースで、もともと死亡などのイベントのリスク低減に対して、劇的な効果を示す治療法はほとんどないため、そこから大きく乖離する結果が得られた場合には、研究者は懐疑的になるべきだ」
- ▶ Peto *et al.* (1995) も「3分の1もの死亡率の低減効果は、非現実的に大きいと考えるべきだ」と同様の意見を述べている
- ▶ ベイズ流の解析をするのであれば、ISIS-4以前のメタアナリシスの時点で、懐疑的事前分布を使った解析を行うべきだった
- ▶ 頻度論的な解析と結論はどのように変わった??

12

Notations

- ▶ Y_1, Y_2, \dots, Y_N : 試験ごとの対数オッズ比の推定量
- ▶ $\sigma_1^2, \sigma_2^2, \dots, \sigma_N^2$: 対数オッズ比の分散の推定値

$$Y_i = \log \left\{ \frac{X_{i1}(n_{i0} - X_{i0})}{(n_{i1} - X_{i1})X_{i0}} \right\},$$
$$\sigma_i^2 = \frac{1}{X_{i1}} + \frac{1}{X_{i0}} + \frac{1}{n_{i1} - X_{i1}} + \frac{1}{n_{i0} - X_{i0}}$$

- ▶ $X_{i1} \sim \text{Binomial}(n_{i1}, p_{i1})$
- ▶ $X_{i0} \sim \text{Binomial}(n_{i0}, p_{i0})$

13

固定効果モデル Fixed-effect model

- ▶ すべての試験における対数オッズ比の真値が共通であると仮定をして、その共通オッズ比の推測を行う
- ▶ 頻度論の解析方法では、条件付き最尤法, Mantel-Haenszel法, PetoのOne-step法, 逆分散法などで、比較的簡単に分析を行うことができる

14

変量効果モデル Random-effects model

- ▶ 試験ごとの試験デザイン, 患者背景, 治療, アウトカムなど、さまざまな要因の異質性によって、治療効果の大きさ (effect size) は異なると想定される
- ▶ 試験ごとの対数オッズ比の真値 $\theta_1, \theta_2, \dots, \theta_N$ は異なると仮定するべき??
- ▶ パラメータ $\theta_1, \theta_2, \dots, \theta_N$ を特定の確率分布 $G(\theta; \xi)$ に従う確率変数 (変量効果; random effects) と仮定する
- ▶ $\theta_1, \theta_2, \dots, \theta_N$ は「未知かつ異なる値をとり得る」という仮定のもと、 $G(\theta; \xi)$ の超パラメータ ξ で “Estimand” を定義し、その推測を行う

15

DerSimonian-Laird式の変量効果モデル

- ▶ 変量効果によって、試験間の異質性をモデル化したモデル
 - ▶ $Y_i \sim N(\theta_i, \sigma_i^2)$ ($i = 1, 2, \dots, N$)
 - ▶ $\theta_i \sim N(\mu, \tau^2)$
- ▶ $\theta_1, \theta_2, \dots, \theta_N$ は、変量効果 (確率変数) であり、正規分布に従うと仮定
- ▶ $\theta_1, \theta_2, \dots, \theta_N$ は未知としたまま、変量効果の分布のパラメータ (こちらが “Estimand” に相当) の推定を行う
 - ▶ μ : 異質性があるもとでの平均的な治療効果
 - ▶ τ^2 : 異質性の大きさを表すパラメータ

16

最尤法

- ▶ $\theta_1, \theta_2, \dots, \theta_N$ を積分消去した Y_1, Y_2, \dots, Y_N の周辺分布から尤度 (周辺尤度) を構成する

$$p(y_i) = \int p(y_i|\theta_i)p(\theta_i|\mu, \tau^2)d\theta_i = \phi(y_i|\mu, \sigma_i^2 + \tau^2)$$

- ▶ 正規分布-正規分布の組み合わせであれば、この積分は簡単に解くことができ、周辺分布も正規分布になるので、最尤法での推測は簡単にできる
 - ▶ $Y_i \sim N(\mu, \sigma_i^2 + \tau^2)$

17

Table 2 Summary data from 15 randomized trials of intravenous magnesium for acute myocardial infarction (data from Sterne *et al.*²⁶) with classical meta-analysis using the Peto method²⁷ and DerSimonian and Laird (D-L) method²⁸ based on simple log(odds ratios) and a moment estimate of between-trial variance. A test of homogeneity of odds ratios (ρ_{het}) is as described by Wood¹⁴

Trial	Magnesium group		Control group	
	Deaths r_i^M	Patients n_i^M	Deaths r_i^C	Patients n_i^C
Morton	1	40	2	36
Rasmussen	9	135	23	135
Smith	2	200	7	200
Abraham	1	48	1	46
Feldtsch	10	150	8	148
Shechter 1990	1	59	9	56
Ceremnevski	1	25	3	23
LMIT-2	90	1150	118	1157
Fixed effect (Peto) meta-analysis of above eight trials: OR = 0.65 (95% CI: 0.51, 0.82); $\rho_{het} = 0.19$				
Random effects (D-L) meta-analysis of above eight trials: OR = 0.55 (95% CI: 0.34, 0.89)				
Berthiaume	0	22	1	21
Singh	6	76	11	75
Pereira	1	27	7	27
Goff	5	23	13	33
Thompson	4	120	8	122
Shahane 1995	4	107	17	108
Fixed effect (Peto) meta-analysis of above 14 trials: OR = 0.57 (95% CI: 0.46, 0.71); $\rho_{het} = 0.19$				
Random effects (D-L) meta-analysis of above 14 trials: OR = 0.47 (95% CI: 0.33, 0.68)				
SSS = 22.0 (21.0)				
Fixed effect (Peto) meta-analysis of above 15 trials: OR = 1.01 (95% CI: 0.95, 1.07); $\rho_{het} = 0.0002$				
Random effects (D-L) meta-analysis of above 15 trials: OR = 0.53 (95% CI: 0.36, 0.77)				

Higgins and Spiegelhalter (2002)

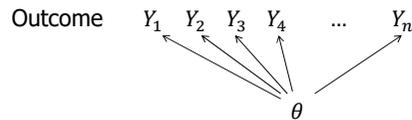
Petoの方法が、条件付き最尤法を近似した固定効果モデルによる方法

異質性を考慮した変量効果モデルによる方法が“D-L”法

いずれも、有意な効果が示唆されている!!

18

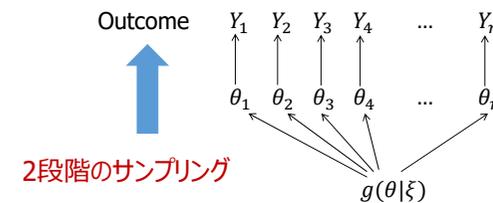
均質性を仮定したモデル



θ を Estimandとする共通の分布から、個々のデータのサンプリングは行われていると仮定

19

異質性を仮定した“マルチレベル”モデル



個々のアウトカム Y_k は、異なる効果 θ_k を持つ異なる分布からサンプリングされていると仮定。
さらに、 θ_k は特定の分布 $g(\theta|\xi)$ から生成されていると仮定。

20

階層モデル Hierarchical model

- ▶ 階層モデル (hierarchical model)
 - ▶ $Y_i \sim f(y|\theta_i), (i = 1, 2, \dots, n)$
 - ▶ $\theta_i \sim g(\theta|\xi)$
 - ▶ $\xi \sim h(\xi)$ 超パラメータ ξ に対する事前分布
- ▶ ベイズ統計学での“階層モデル”とは、頻度論でも使われるマルチレベルモデルの(超)パラメータに事前分布を仮定して、ベイズモデルとしたモデルのこと
- ▶ 複数レベルの階層構造を持ったモデルなので、このように呼ばれる

21

Binomial-Normal型のマルチレベルモデル

- ▶ 正規-正規型のモデルでは、 σ_i^2 を既知としていたが、推定値をPlug-inするような場合、誤差を無視していることになる
- ▶ 2項分布から直接、階層モデルを構築することもできる
- ▶ $X_{i1} \sim \text{Binomial}(n_{i1}, p_{i1})$
- ▶ $X_{i0} \sim \text{Binomial}(n_{i0}, p_{i0})$
- ▶ $\theta_i \sim N(\mu, \tau^2), (i = 1, 2, \dots, n)$ $\theta_i = \log \left\{ \frac{p_{i1}(1-p_{i0})}{(1-p_{i1})p_{i0}} \right\}$

22

周辺尤度の計算

- ▶ $\theta_1, \theta_2, \dots, \theta_N$ を積分消去した周辺分布から周辺尤度を作ると...

$$p(x_{ij}) = \int p_{ij}^{x_{ij}} (1-p_{ij})^{1-x_{ij}} \phi(\theta_i|\mu, \tau^2) d\theta_i$$

$(i = 1, 2, \dots, n; j = 0, 1)$

- ▶ 周辺尤度の積分が...解けない??
- ▶ ベイズ統計で使う共役モデルのように、異なるレベルの確率分布モデルの相性が良くない場合、この積分は解くことができず、周辺尤度に基づく最尤法での推測は一般的に困難!!

厳密には、先ほどのモデルと尤度が一致しませんが、ご容赦ください。

23

より一般的な定式化

- ▶ 結果変数 $\mathbf{Y}_i = (Y_{i1}, \dots, Y_{in_i})^T (i = 1, 2, \dots, N)$
 - ▶ 指数型分布 (正規分布, 2項分布, ポアソン分布など) に従うものとする
- ▶ 共変量行列 $\mathbf{X}_i = (x_{i1}^T, \dots, x_{in_i}^T)^T, \mathbf{Z}_i = (z_{i1}^T, \dots, z_{in_i}^T)^T$
- ▶ 回帰関数 $\mu_i = E(\mathbf{y}_i|\mathbf{X}_i, \mathbf{Z}_i)$

24

一般化線形混合効果モデル

- ▶ Generalized Linear Mixed Model (GLMM)
 - ▶ $g\{E(y_i|X_i, Z_i, \eta_i)\} = X_i\beta + Z_i\eta_i$
 - ▶ 変量効果 η_i が与えられたもとで、 $\{y_{it}\}$ が条件付き独立となるという仮定を置く
 - ▶ η_i はなんらかの確率分布（一般的には、正規分布）に従う変量効果
- ▶ $g(\cdot)$: 適当なリンク関数
- ▶ X_i : 個人ごとの固定効果に対応する共変量行列
- ▶ Z_i : 個人ごとの変量効果に対応する共変量行列

Diggle et al. (2002), 高井ら (2016) 25

線形混合効果モデル

- ▶ $y_i = X_i\beta + Z_i\eta_i + \varepsilon_i$
 - ▶ $\eta_i \sim \text{MVN}(0, S)$, $\varepsilon_i \sim \text{MVN}(0, \Sigma)$
- ▶ y_i の分布型が正規分布、リンク関数が恒等リンク (identity link) であり、変量効果の分布が正規分布であるモデル
- ▶ 近年、医薬品開発の臨床試験で広く用いられるようになった MMRM (mixed effect models for repeated measurements) は、このモデルの亜型 (変型)

Diggle et al. (2002), Mallinckrodt et al. (2001) 26

GLMMにおける最尤法

- ▶ 一般的に、GLMMのもとでは、積分計算を解析的に解くことができないため、正確な周辺尤度が簡単には求まらない
- ▶ ロジスティック混合効果モデル、ポアソン混合効果モデルでさえ、周辺尤度は正確に計算することができない！
- ▶ 数値積分やラプラス近似などに基づく近似周辺尤度による、近似的な最尤法を行うことが一般的である
- ▶ 正規分布を仮定した線形混合効果モデルは例外で、この積分計算が可能であり、1980年代頃から、経時データの解析への応用に用いられてきた (Laird and Ware, 1982)

27

Statistics
in Medicine

Research Article

On fitting generalized linear mixed-effects models for binary responses using different statistical packages

Hui Zhang, Naiji Lu, Changyong Feng, Sally W. Thurston, Yinglin Xia, Liang Zhu, Xin M. Tu

First published: 10 June 2011 | <https://doi.org/10.1002/sim.4265> | Citations: 65

Read the full text > PDF TOOLS SHARE

Abstract

The generalized linear mixed-effects model (GLMM) is a popular paradigm to extend models for cross-sectional data to a longitudinal setting. When applied to modeling binary responses, different software packages and even different procedures within a package may give quite different results. In this report, we describe the statistical approaches that underlie these different procedures and discuss their strengths and weaknesses when applied to fit correlated binary responses. We then illustrate these considerations by applying these procedures implemented in some popular software packages to simulated and real study data. Our simulation results indicate a lack of reliability for most of the procedures considered, which carries significant implications for applying such popular software packages in practice. Copyright © 2011 John Wiley & Sons, Ltd.

<https://onlinelibrary.wiley.com/doi/10.1002/sim.4265>

28

Table III. Estimates of parameters, standard errors, and type I errors for the simulation model with $n=500$ using different procedures/packages.

τ	Comparison of estimates across procedures/packages for the simulation model with $n=500$								
	Software	Procedure/package	Type I error	$\beta_0=1$	s.e. $\times 10$	$\beta_1=1$	s.e. $\times 10$		
0.001	SAS	NLMIXED	0.046	1.025	0.75	1.029	0.85		
		GLIMMIX	REML	0.034	1.004	0.67	1.004	0.75	
			ML	0.036	1.005	0.67	1.003	0.75	
			lme4	Laplace	0.094	1.029	0.67	1.036	0.75
				Gauss-Hermite	0.098	1.022	0.67	1.029	0.75
			ZELIG	Laplace	0.094	1.029	0.67	1.036	0.75
	R		Gauss-Hermite	0.097	1.022	0.67	1.029	0.75	
		glmml	Laplace	0.040	1.014	0.69	1.013	0.76	
			Gauss-Hermite	0.044	1.016	0.69	1.015	0.77	
		SAS	NLMIXED	0.049	1.003	0.69	1.002	0.82	
			GLIMMIX	REML	0.192	0.912	0.66	0.897	0.75
				ML	0.162	0.914	0.67	0.902	0.77
	lme4		Laplace	0.117	1.010	0.70	1.013	0.82	
			Gauss-Hermite	0.118	1.003	0.70	1.001	0.81	
	ZELIG		Laplace	0.117	1.010	0.70	1.013	0.82	
R		Gauss-Hermite	0.117	1.003	0.70	1.001	0.81		
	glmml	Laplace	0.152	0.977	0.76	0.932	0.80		
		Gauss-Hermite	0.126	0.988	0.78	0.943	0.82		
	SAS	NLMIXED	0.066	0.992	1.25	0.976	1.66		
		GLIMMIX	REML	0.992	0.665	0.88	0.519	1.05	
			ML	0.993	0.586	0.87	0.467	1.01	
		lme4	Laplace	0.249	1.056	1.29	1.114	1.70	
			Gauss-Hermite	0.256	0.983	1.23	0.912	1.41	
		ZELIG	Laplace	0.247	1.056	1.29	1.113	1.70	
R		Gauss-Hermite	0.209	0.988	1.23	0.955	1.49		
	glmml	Laplace	0.416	0.962	1.42	0.767	1.42		
		Gauss-Hermite	0.385	0.953	1.52	0.758	1.53		

SAS, Rなどで、当時利用することができたGLMMの標準的なパッケージはすべて、妥当な推測ができなかったという衝撃のシミュレーション結果が！！(ロジスティック混合効果モデル)

現在まで、この問題が解決できているかは、わかりません(が、解決できているのでしょうか！)

ここまでのまとめ

- ▶ 一般化線形混合モデルを含め、マルチレベルモデルでの分析は、現代応用統計学における重要な方法
- ▶ しかし、頻度論での分析は、尤度関数が解析的に求められないという大きな計算上の問題がある(近似計算にも、精度上の問題があるという説もある)
- ▶ 一方、ベイズ流の階層モデルとして定式化をしてやれば...
- ▶ どのような複雑な階層構造を持つ場合も、非共役な分布の組み合わせでも、MCMCを使えば、統一的な枠組みのもとで、事後推測ができる！！

Binomial-Normal型のマルチレベルモデル

- ▶ $X_{i1} \sim \text{Binomial}(n_{i1}, p_{i1})$
- ▶ $X_{i0} \sim \text{Binomial}(n_{i0}, p_{i0})$
- ▶ $\theta_i \sim N(\mu, \tau^2), (i = 1, 2, \dots, n)$
- ▶ $\{\mu, \tau^2\}$ に、事前分布を仮定しよう！
- ▶ Higgins and Spiegelhalter (2002) は、以下の無情報事前分布を考えた (reference Bayesian approach)
 - ▶ $\mu \sim N(0.0, 10^5)$
 - ▶ $\tau \sim U(0, 100)$

MCMCのための汎用的なソフトウェア。WinBUGSの開発・サポートは終了しており、OpenBUGSに移行している。

R2OpenBUGS: Running OpenBUGS from R
Using this package, it is possible to call a BUGS model, summarize inferences and convergence in a table and graph, and save the simulations in arrays for easy access in R.

Version: 3.2-3.2.1
 Depends: R (≥ 2.13.0)
 Imports: coda (≥ 0.11-0), boot
 Published: 2020-04-02
 Author: originally written as R2WinBUGS by Andrew Gelman; changes and packaged by Gábor Szturc and Uwe Ligges. With considerable contributions by Gregor Gorjanc and Jouni Karman. Adapted to R2OpenBUGS from R2WinBUGS by Neal Thomas.
 Maintainer: Neal Thomas <nthomas@u.washington.edu>

License: GPL-2
 NeedsCompilation: no
 SystemRequirements: OpenBUGS (≥ 3.2.2)
 Citation: R2OpenBUGS citation info
 Materials: NEWS
 In review: CRAN@rhub.io, RnewsModels
 CRAN checks: R2OpenBUGS results

Recent test:
 Reference manual: R2OpenBUGS.pdf
 Vignettes: R2OpenBUGS

Source code: R2OpenBUGS 3.2-3.2.1.tar.gz
 Windows binaries: r-devel: R2OpenBUGS_3.2-3.2.1.zip, r-release: R2OpenBUGS_3.2-3.2.1.zip, r-oldrel: R2OpenBUGS_3.2-3.2.1.zip
 macOS binaries: r-release (arm64): R2OpenBUGS_3.2-3.2.1.tar, r-oldrel (arm64): R2OpenBUGS_3.2-3.2.1.tar, r-release (x86_64): R2OpenBUGS_3.2-3.2.1.zip, r-oldrel (x86_64): R2OpenBUGS_3.2-3.2.1.zip
 Old sources: R2OpenBUGS archive

Reverse dependencies:
 Reverse imports: coda
 Reverse suggests: AICcmodexp, ISIsibs, postback

Use the canonical form <https://doi.org/10.18122/cran.r-project.org/web/packages/R2OpenBUGS/> to link to this page.

<https://cran.r-project.org/web/packages/R2OpenBUGS/>

Rから、OpenBUGSを操作するためのパッケージ。Andrew Gelmanらが開発したR2WinBUGSをOpenBUGS用に改変したもの。

OpenBUGSは、完全にバックグラウンドで動かすことができ、BUGSのモデルコードも、すべてR上で定義することができる。

以前のR2WinBUGSに比べると、かなり高速で、動作もスムーズになっている。ただし、大規模データを扱う場合は、かなり計算コストがかかる。大規模データの解析は、Stanのほうが良い。

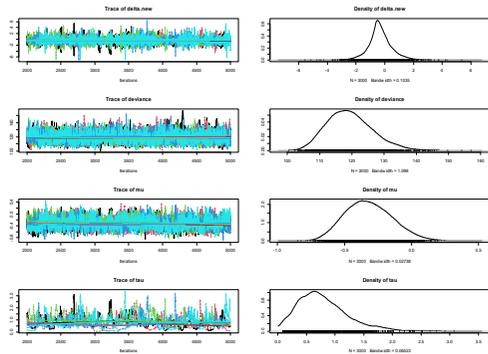
OpenBUGSによるMCMC

```
rma01 <- function(){
  for (i in 1:N) {
    rc[i] ~ dbin(p[i],nc[i]);
    rm[i] ~ dbin(p[i],nm[i]);
    phi[i] <- logit(p[i]);
    logit(p[i]) <- phi[i] + delta[i];
    delta[i] ~ dnorm(mu, precision);
    p[i] ~ dunif(0,1);
  }
  delta.new ~ dnorm(mu, precision);
  mu ~ dnorm(0, 0.0001);
  # Reference prior for mu
  # mu ~ dnorm(0, 12.45);
  tau ~ dunif(0, 100);
  precision <- 1/(tau*tau);
  tau.sq <- 1/precision;
  OR <- exp(mu);
  leas0 <- min(delta.new, 0);
  prob0 <- 1 - equals(leas0, 0);
  leas10 <- min(delta.new, -0.1084);
  prob10 <- 1 - equals(leas10, -0.1084);
}
```

```
mainite <- function(){
  list(mu = 0, tau = 1)
}
rmaout.01 <- bugs(data=mdatat4,
  inite=mainite,
  parameters.to.save=c("OR","mu","tau","delta.new","prob0","prob10"),
  model.file=mainit,
  n.chains=1,
  n.iter=1000,
  n.burnin=500)
rmaout.01$summary
```

Higgins and Spiegelhalter (2002) に公開されているBUGSコードをR2OpenBUGSで動かすための事例プログラムを、以下のURLで公開しています。

http://www.ism.ac.jp/~noma/ex_BUGS.r



codaのツールを使うことによって、R2OpenBUGSの出力から、MCMCの収束診断などの分析を行うことができます。

R2OpenBUGSによるMCMC

```
> rmaout.01$summary
              mean      sd      2.5%      25%      50%      75%      97.5%
OR           0.3665278 0.09589621 0.1965950 0.299700 0.3612 0.427000 0.5666025
mu          -1.0392638 0.27225207 -1.6270000 -1.205000 -1.0180 -0.850975 -0.5680000
tau           0.6085564 0.26880441 0.2130975 0.421075 0.5669 0.751500 1.2460000
delta.new   -1.0467332 0.71800536 -2.5920000 -1.437000 -1.0080 -0.624600 0.3320025
prob0       0.9471000 0.22383944 0.0000000 1.0000000 1.0000 1.000000 1.0000000
prob10      0.9301000 0.25498479 0.0000000 1.0000000 1.0000 1.000000 1.0000000
deviance    117.2297795 6.56044651 106.2000000 112.600000 116.6000 121.300000 131.8000000
```

どのような複雑な階層モデルを使った場合も、ベイズ解析では、MCMCを使って統一的に解析を行うことができる。Flat Priorを使えば、頻度論による分析を近似した推測を行うこともできる。もちろん、積極的な事前情報を組み込んだベイズ流の解析を行うこともできる。

懐疑的事前分布の設定について

- ▶ We base a sceptical *a priori* position on Pogue and Yusuf: “Most clinically important interventions are likely to reduce the relative risk of major outcomes, such as myocardial infarction, stroke, or death, by about 10-20%”. We might then consider that a reasonable degree of scepticism is to think it unlikely (only 5% chance) that magnesium would reduce the odds of mortality by more than 25%. In section (2) of Appendix we show how this can be translated into a normal prior distribution for the common log (odds ratio), centered on 0 and with a variance of 0.03, and that this precision is equivalent to a “trial” with 72 deaths in each group.
- ▶ 事前の懐疑的な意見をもとに、マグネシウムが死亡率のオッズを25%低減する確率を5%と設定。これに対応する平均 0 の正規分布は、分散 0.03 の分布になることに。これを事前分布に設定することに。これは、データに換算すると、群ごとに72例の死亡が起こった試験と等価な情報を持つ事前分布となる。

Higgins and Spiegelhalter (2002) 37

懐疑的事前分布によるベイズ解析の結果

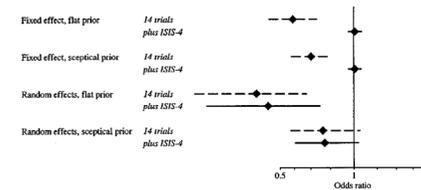


Figure 3 The impact of adding ISIS-4 to different meta-analyses of the 14 smaller studies (including LIMIT-2). 95% posterior credible intervals for the mortality odds ratio associated with magnesium, for both fixed effect and random effects analysis, with a ‘flat’ reference prior and the reasonably sceptical prior (5% chance of at least a 25% reduction in mortality odds)

Higgins and Spiegelhalter (2002)

懐疑的事前分布を用いたベイズ流の解析では、頻度論的な解析から、一転して、95%CrIは1を跨ぐという結果に！

ISIS-4による否定的なエビデンスが出る以前にも、慎重な結論を出すべきという結果が得られている

38

交換可能性 Exchangeability

- ▶ 階層ベイズモデル
 - ▶ $Y_i \sim f(y|\theta_i)$, ($i = 1, 2, \dots, n$) データの分布
 - ▶ $\theta_1, \theta_2, \dots, \theta_n \sim g(\theta|\xi)$ 変量効果の分布
 - ▶ $\xi \sim h(\xi)$ 事前分布
- ▶ 試験ごとのパラメータ $\theta_1, \theta_2, \dots, \theta_n$ が同一分布 ($g(\theta|\xi)$) に従うという仮定が、階層モデルでの分析を行う上での肝！！
- ▶ ベイズ統計の文献では、この仮定は、“i.i.d.”ではなく、交換可能 (exchangeable) と書かれていることが多い

39

交換可能性とは？

- ▶ 交換可能性：文献上の定義は...
 - ▶ $(\theta_1, \theta_2, \dots, \theta_n)$ の要素を任意の順序に並び替えたとしても、その同時分布は不変であるという概念
- ▶ de Finettiの定理などから誘導される概念で、i.i.d. よりも少し緩い理論的仮定となる (i.i.d. であれば、必ず交換可能)
- ▶ しかし、「i.i.d.ではないが、交換可能」という例は、かなりイレギュラーなものばかりであり (cf. Pólyaの壺モデル)、医学研究の応用で使われるレギュラーなモデルでは、実質的に、i.i.d.を仮定していると考えても差し支えない

40

In Medicine
Research Article
A Bayesian hierarchical survival model for the institutional effects in a multi-centre cancer clinical trial
Yutaka Matsuyama, Junichi Sakamoto, Yasuo Ohashi
First published: 21 December 1998 | [https://doi.org/10.1002/\(SICI\)1097-0258\(199809\)15:17:17-1833::AID-SIM878>3.0.CO;2-R](https://doi.org/10.1002/(SICI)1097-0258(199809)15:17:17-1833::AID-SIM878>3.0.CO;2-R) | Citations: 21

Abstract
In randomized clinical trials comparing treatment effects on diseases such as cancer, a multi-centre trial is usually conducted to accrue the required number of patients in a reasonable period of time. While we interpret the average treatment effect, it is necessary to examine the homogeneity of the observed treatment effects across institutions, that is, treatment-by-institution interaction. If the homogeneity is confirmed, the conclusions concerning treatment effects can be generalized to a broader patient population; in this paper, a Bayesian hierarchical survival model is used to investigate the institutional effects on the efficacy of treatment as well as on the baseline risk. The marginal posterior distributions are estimated by a Markov chain Monte Carlo method, that is, Gibbs sampling, to overcome current computational limitations. The robustness of the inferences to the distributional assumption for the random effects is also examined. We illustrate the methods with analyses of data from a multi-centre cancer clinical trial, which investigated the efficacy of irinotecan/5-fluorouracil as an adjuvant treatment after curative resection of gastric cancer. In this trial there is little difference in the treatment effects across institutions and the treatment is shown to be effective, while there appears to be substantial variations in the baseline risk across institutions. This result indicates that the observed treatment effects might be generalized to a broader patient population. © 1998 John Wiley & Sons, Ltd.

Recommended
Bayesian Hierarchical Model
Practical Applications of Bayesian Reliability (1)
Bayesian model selection using the median probability model
Type Check
Wiley Computational Statistics

[https://doi.org/10.1002/\(SICI\)1097-0258\(199809\)15:17:17-1833::AID-SIM878>3.0.CO;2-R](https://doi.org/10.1002/(SICI)1097-0258(199809)15:17:17-1833::AID-SIM878>3.0.CO;2-R)

41

SIP Study

- ▶ 胃がん治癒切除後の免疫化学療法の有効性を評価した多施設臨床試験 (N=253)
- ▶ 施設内の患者数が2名以上の36施設が対象
- ▶ 標準的な補助化学療法を行う群と、標準治療に加えて免疫療法を行う群を比較したランダム化比較試験
- ▶ 死亡までの生存時間がエンドポイント
- ▶ Matsuyama et al. (1998) では、施設間の異質性を評価するために、階層モデルを用いた解析を行っている

Nakazato et al. (1994) 42

階層比例ハザード回帰モデル

- ▶ 階層比例ハザード回帰モデル

$$\lambda_{ij}(t|x_{ijk}, b_{i0}, b_{i1}) = \lambda_0(t) \exp\left(\sum_{k=1}^p x_{ijk}\beta_k + b_{i0} + b_{i1}x_{ij1}\right)$$

$$\begin{pmatrix} b_{i0} \\ b_{i1} \end{pmatrix} \sim N\left[\begin{pmatrix} 0 \\ 0 \end{pmatrix}, \begin{pmatrix} \theta_{11} & \theta_{12} \\ \theta_{12} & \theta_{22} \end{pmatrix}\right]$$

- ▶ $\lambda_{ij}(t)$: i 番目の施設の j 番目の患者のハザード関数
- ▶ x_{ijk} : i 番目の施設の j 番目の患者の k 番目の共変量 (ただし、 $k=1$ は治療を表す指示変数)
- ▶ β_k : k 番目の共変量の固定効果パラメータ
- ▶ b_{i0}, b_{i1} : 施設の効果, 治療-施設の交互作用を表す変量効果

Matsuyama et al. (1998) 43

Gibbs Sampling

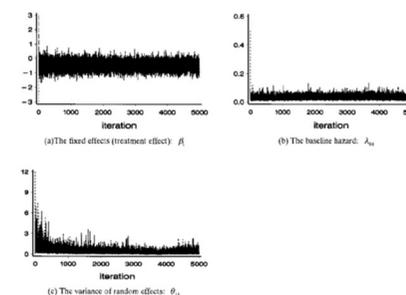


Figure 2. The traces of the sampled values from three independent runs of the Gibbs sampling applied to model (1)

無情報事前分布を仮定したもとのMCMC (Matsuyama et al., 1998)

44

MCMCの結果

Matsuyama et al. (1998)

Table II. Posterior summaries for the model parameters (Normal distribution)

Parameter	Mean	Median	Standard deviation	5 percentile	95 percentile	\sqrt{R}
Treatment	-0.472	-0.471	0.295	-0.951	0.011	1.00
p_{TNM1}	-1.849	-1.847	0.630	-2.890	-0.830	1.00
p_{TNM2}	-0.298	-0.303	0.334	-0.831	0.270	1.00
p_{TNM3}	1.657	1.644	0.300	1.186	2.173	1.00
p_{TNM4}	2.439	2.433	0.473	1.664	3.219	1.00
$\phi_{treatment}$	0.652	0.624	0.197	0.386	1.011	1.00
λ_{01}	0.035	0.033	0.014	0.016	0.062	1.00
λ_{02}	0.137	0.131	0.043	0.075	0.214	1.00
λ_{03}	0.064	0.061	0.026	0.030	0.111	1.00
λ_{04}	0.070	0.066	0.029	0.033	0.124	1.00
λ_{05}	0.079	0.074	0.032	0.036	0.139	1.00
λ_{06}	0.050	0.044	0.028	0.015	0.105	1.00
θ_{11}	0.340	0.280	0.253	0.057	0.826	1.01
θ_{12}	-0.010	-0.001	0.094	-0.161	0.109	1.00
θ_{22}	0.100	0.050	0.141	0.001	0.368	1.00

* θ_{kl} denotes the (k, l) element of the covariance matrix for the random effects

cf.) 施設で層別した層別Cox回帰のHR: 0.55 (90%CI: 0.35-0.86)

45

施設の効果の事後推測

Matsuyama et al. (1998)

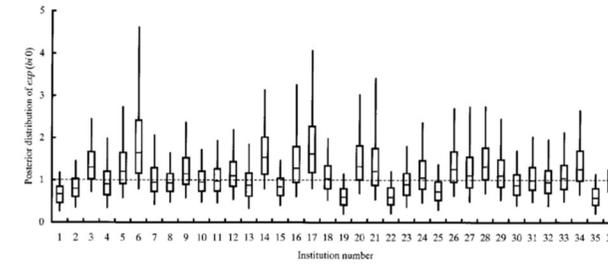


Figure 3. Posterior distribution of the baseline hazard in each institution

施設間の異質性はかなり大きく、施設が強い予後因子になることが示唆される

46

治療-施設の交互作用

Matsuyama et al. (1998)

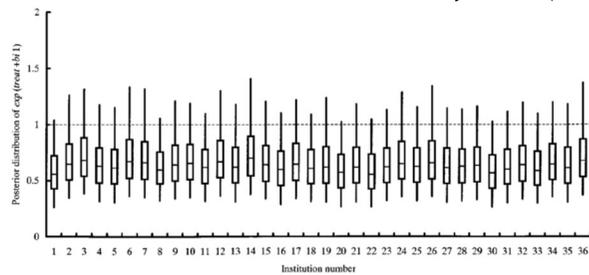


Figure 4. Posterior distribution of the treatment effects in each institution

施設ごとのハザード比の推定値 (それほど大きな施設間差はない)

47

Statistics in Medicine

Original

Bayesian subset analysis in a colorectal cancer clinical trial

Dennis O. Dixon, Richard Simon

First published: 1992 | <https://doi.org/10.1002/sim.4780110104> | Citations: 51

PDF TOOLS SHARE

References Related Information

Recommended

Bayesian Analysis
Big Data and Differential Privacy: Analysis Strategies for Railway Track Engineering, [1]

Bayesian Reliability Analysis
Harry F. Martz
Wiley StatsRef: Statistics Reference Online, [1]

Bayesian Reliability Analysis
Harry F. Martz
Encyclopedia of Statistics in Quality and Reliability, [1]

<https://onlinelibrary.wiley.com/doi/abs/10.1002/sim.4780110104>

48

臨床試験のサブグループ解析への応用

- ▶ さまざまな要因でサブグループ解析を行うと、“By chance”で、極端に大きな差が出てしまったり、有意差が出てしまうこともある（検定の多重性の問題）
- ▶ Dixon and Simon (1991, 1992) は、階層ベイズモデルを使って、サブグループごとの治療効果を縮小推定する分析方法を提案した
- ▶ Bayesian subgroup analysisは、米国の医療機器の臨床試験でも、近年、用いられている（Campbell et al., 2023; ここで紹介する方法をさらに高度化したもの）

49

階層比例ハザード回帰モデル

- ▶ 対象となる共変量の交互作用項をすべて含めたCox回帰モデル

$$\lambda(t|z, \mathbf{x}) = \lambda_0(t) \exp \left(z\tau + \sum_{i=1}^m x_i(\beta_i + z\gamma_i) \right)$$

$$\begin{aligned} p(\tau) &\propto 1, p(\beta_i) \propto 1 \\ p(\gamma_i) &\propto N(0, \xi^2) \\ p(\xi^2) &= [\max(\xi^2, 0.005)]^{-1} \end{aligned}$$

- ▶ z : treatment indicator, x_i : covariate

Dixon and Simon (1992)

サブグループ効果を表す交互作用パラメータ $\gamma_1, \dots, \gamma_m$ に交換可能性を仮定し、Ridge回帰と似たような縮小推定を行うことを狙った階層モデル

50

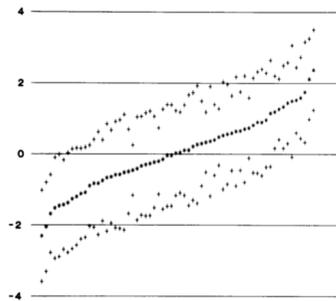


Figure 3. Subnet-specific treatment effect estimates and 95 per cent confidence limits using regression results. (+) Point estimate (+) 95 per cent confidence limit

Dixon and Simon (1992)

51

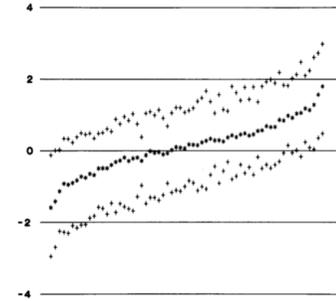


Figure 4. Subnet-specific treatment effect point and interval estimates using the Bayesian method. (+) Posterior mode (+) 95 per cent highest posterior density interval limit

Dixon and Simon (1992)

52

Table III. Point and interval estimates of standardized treatment effects

Category	Regression estimates		Bayesian estimates	
	Point	Interval	Point	Interval
Performance status				
0-1	0.36	[- 0.07, 0.78]	0.31	[- 0.09, 0.72]
2-3	- 0.27	[- 1.31, 0.77]	- 0.07	[- 0.98, 0.70]
Anaplasia grade				
1-2	0.07	[- 0.37, 0.50]	0.12	[- 0.30, 0.53]
3-4	0.92	[0.10, 1.74]	0.66	[- 0.04, 1.44]
Measurable disease				
No	0.78	[0.21, 1.35]	0.62	[0.07, 1.19]
Yes	- 0.16	[- 0.68, 0.37]	- 0.05	[- 0.56, 0.44]
Symptomatic				
No	0.70	[0.08, 1.33]	0.57	[- 0.01, 1.18]
Yes	- 0.08	[- 0.57, 0.40]	- 0.01	[- 0.48, 0.45]
Age				
< 70 years	0.52	[0.09, 0.95]	0.42	[0.00, 0.85]
≥ 70 years	- 0.70	[- 1.61, 0.20]	- 0.38	[- 1.23, 0.39]
Sex				
Male	0.16	[- 0.33, 0.65]	0.17	[- 0.29, 0.62]
Female	0.41	[- 0.22, 1.04]	0.35	[- 0.21, 0.94]

Dixon and Simon (1992)

53

FDAの医療機器臨床試験のガイダンス

Guidance for the Use of Bayesian Statistics in Medical Device Clinical Trials

The screenshot shows the official FDA guidance document page. It includes the title 'Guidance for the Use of Bayesian Statistics in Medical Device Clinical Trials', the date 'FEBRUARY 2010', and the issuing office 'Office of the Commissioner, Office of Clinical Policy and Programs'. It also provides contact information for Dr. Greg Campbell and the document's issuance date of February 5, 2010. The page features navigation links for PDF, HTML, and Print, and a search bar for FDA guidance documents.

<https://www.fda.gov/regulatory-information/search-fda-guidance-documents/guidance-use-bayesian-statistics-medical-device-clinical-trials>

54

SPRINGER LINK

Find a journal Publish with us Search

Home > Therapeutic Innovation & Regulatory Science > Article

Review | Published: 03 March 2023

Bayesian Statistics for Medical Devices: Progress Since 2010

Gregory Campbell¹, Jelba Itony, Gene Perrillo & Laura Thompson

Therapeutic Innovation & Regulatory Science 57, 453–463 (2023) | [Cite this article](#)

1820 Accesses | 1 Citations | Metrics

Abstract

The use of Bayesian statistics to support regulatory evaluation of medical devices began in the late 1990s. We review the literature, focusing on recent developments of Bayesian methods, including hierarchical modeling of studies and subgroups, borrowing strength from prior data, effective sample size, Bayesian adaptive designs, pediatric extrapolation, benefit-risk decision analysis, use of real-world evidence, and diagnostic device evaluation. We illustrate how these developments were utilized in recent medical device evaluations. In Supplementary Material, we provide a list of medical devices for which Bayesian statistics were used to support approval by the US Food and Drug Administration (FDA), including those since 2010, the year the FDA published their guidance on Bayesian statistics for medical devices. We conclude with a discussion of current and future challenges and opportunities for Bayesian statistics, including artificial intelligence/machine learning (AI/ML) Bayesian modeling, uncertainty quantification, Bayesian approaches using propensity scores, and computational challenges for high dimensional data and models.

2010年以降、ベイズ統計による臨床試験でFDAが承認した医療機器がすべてリストアップされており、どのような方法が使われていたかもレビューされている。

<https://link.springer.com/article/10.1007/s43441-022-00495-w>

55

THE LANCET

Submit Article Log in Register Subscribe Claim

ARTICLES | VOLUME 395, ISSUE 10206, P1840-1850, OCTOBER 21, 2023 | [Download Full Issue](#)

The screenshot shows the Lancet article page for 'Ultrathin, bioresorbable polymer sirolimus-eluting stents versus thin, durable polymer everolimus-eluting stents in patients undergoing coronary revascularisation (BIOFLOW V): a randomised trial'. It lists authors Dr David E Kandzari, Prof Laura Mauri, Prof Jacques J Koolen, Prof Joseph M Massaro, Prof George Deros, Prof Hector M Garcia-Garcia, and Prof Victor M Garcia-Garcia. The article was published on August 26, 2017, with a DOI of 10.1016/S0140-6736(17)32249-3. The page includes options to purchase, subscribe, save, share, request, and request institutional access.

Summary

Background

The development of coronary drug-eluting stents has included use of new metal alloys, changes in stent architecture, and use of bioresorbable polymers. Whether these advancements improve clinical safety and efficacy has not been shown in previous randomised trials. We aimed to examine the clinical outcomes of a bioresorbable polymer sirolimus-eluting stent compared with a durable polymer everolimus-eluting stent in a broad patient population undergoing percutaneous coronary intervention.

[https://www.thelancet.com/JOURNALS/LANCET/ARTICLE/PIIS0140-6736\(17\)32249-3/FULLTEXT](https://www.thelancet.com/JOURNALS/LANCET/ARTICLE/PIIS0140-6736(17)32249-3/FULLTEXT)

階層モデルを用いたベイズ流の解析が使用された医療機器の臨床試験の事例として紹介されていた BIOFLOW V試験。

56

BIOFLOW V試験

- ▶ 薬剤溶出型冠動脈ステント (drug-eluting stent; N=884) と別な冠動脈ステント (N=450) の有効性を比較したランダム化比較試験
- ▶ 過去の類似した2つの試験 (BIOFLOW II, VI試験) のデータを事前情報として利用している
- ▶ 試験間の効果の異質性を表す「バイアスパラメータ」を導入し、階層モデルでその不確実性をモデル化している
- ▶ 主要エンドポイントは、12カ月時点での標的病変不全率 (target lesion failure rate) であり、コントロール治療に対しての非劣性を示すことが目標とされた

57

Model

For ease of presentation we assume that no censoring occurs. The below model will accommodate censoring.

For a subject i in the BIOFLOW-V study, if $Y_{X,i}^V$ is the time to event, let

$$Y_{X,i}^V \sim \text{exponential}(\lambda_X^V) \text{ and } Y_{O,i}^V \sim \text{exponential}(\lambda_O^V)$$

be the distributions of the time (years) to event in the Xience and Orsiro groups, respectively. The values λ_X^V and λ_O^V are the hazards in the Xience and Orsiro groups.

Evidence from BIOFLOW-IV study: It is first discounted using the discount factor a^{IV} . The discounted data are then assumed to follow exponential distributions with hazard rates λ_X^{IV} and λ_O^{IV} in the Xience and Orsiro groups, respectively. For a subject i in the BIOFLOW-IV study, if $Y_{X,i}^{IV}$ is the time to event, let

$$Y_{X,i}^{IV} \sim \text{exponential}(\lambda_X^{IV}) \text{ and } Y_{O,i}^{IV} \sim \text{exponential}(\lambda_O^{IV})$$

A bias term is assumed to link the TLF hazard rates in BIOFLOW-V and BIOFLOW-IV in each treatment group

$$\bullet \log(\lambda_X^V) = \log(\lambda_X^{IV}) + \delta_X^V \text{ and } \log(\lambda_O^V) = \log(\lambda_O^{IV}) + \delta_O^V$$

Evidence from BIOFLOW-II study: It is first discounted using the discount factor a^{II} . The discounted data are then assumed to follow Exponential distributions with hazard rates λ_X^{II} and λ_O^{II} in the Xience and Orsiro groups, respectively,

$$Y_{X,i}^{II} \sim \text{exponential}(\lambda_X^{II}) \text{ and } Y_{O,i}^{II} \sim \text{exponential}(\lambda_O^{II})$$

A bias term is assumed to link the TLF rates in BIOFLOW-II and BIOFLOW-V in each treatment group.

$$\log(\lambda_X^V) = \log(\lambda_X^{II}) + \delta_X^V \text{ and } \log(\lambda_O^V) = \log(\lambda_O^{II}) + \delta_O^V$$

https://classic.clinicaltrials.gov/ProvidedDocs/46/NCT02389946/SAP_001.pdf

58

Prior distribution

The bias terms are assumed to follow a normal distribution with mean 0 and standard deviation τ_1 and τ_2 .

$$\delta_X^V \sim \text{Normal}(0, \tau_1) \text{ and } \delta_O^V \sim \text{Normal}(0, \tau_1)$$

$$\delta_X^{IV} \sim \text{Normal}(0, \tau_2) \text{ and } \delta_O^{IV} \sim \text{Normal}(0, \tau_2)$$

As with primary endpoint analysis, a value of 0.3 will be assumed for both τ_1 and τ_2 . Non-informative prior distributions are assumed for TLF rates in the BIOFLOW-V study

$$\lambda_X^V \sim \text{gamma}(0.1, 0.1) \text{ and } \lambda_O^V \sim \text{gamma}(0.1, 0.1)$$

The non-inferiority condition will be expressed in terms of 1-year TLF rates in the BIOFLOW-V study as with the primary hypothesis

$$P(H_A | \text{Data}) = P(\pi_X^V - \pi_O^V > -\delta | \text{Data}) > \pi^*$$

where the 12-month TLF rates will be calculated as

$$\pi_X^V = 1 - \exp(-\lambda_X^V) \text{ and } \pi_O^V = 1 - \exp(-\lambda_O^V)$$

BIOFLOW II, IV試験の個人レベルのデータを、事前分布に要約せずに「統合する」という解析なので、どちらかというと、個人データメタアナリシスに近い。

https://classic.clinicaltrials.gov/ProvidedDocs/46/NCT02389946/SAP_001.pdf

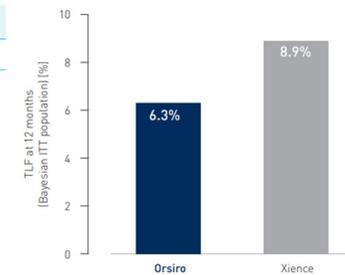
59

Primary endpoint at 12 months (Bayesian ITT population)

	Orsiro n = 1,466	Xience n = 742
TLF, posterior mean	6.3 ± 0.8	8.9 ± 1.2

Estimated effect size at 12 months

-2.6% TLF



http://biotronik.cdn.mediamid.com/cdn_bio_doc/bio26904/26419/bio26904.pdf

60

Bayesian posterior probability on TLF at 12 months

Non-inferiority



Interpretation

There is 100% confidence that the difference in TLF at 12 months is smaller than 3.85% [pre-defined non-inferiority margin] demonstrating unequivocal non-inferiority. The estimated treatment effect of Orsiro is a reduction of TLF by absolute 2.6%.

http://biotronik.cdn.mediamid.com/cdn_bio_doc/bio26904/26419/bio26904.pdf

61

おわりに

- ▶ 階層モデル (+MCMC) を利用した解析によって、従来の頻度論の枠組みでは到底不可能であった、複雑かつ柔軟なモデルを用いた分析が可能となる
- ▶ また、事前情報を活用したベイズ流の解析を主解析に据えた臨床試験も、既に、実践上、多く行われるようになっており、医療機器開発を中心に、ひとつのスタンダードとなっている
- ▶ 今後のベイズ解析の実践において、階層モデルによるモデリングは必須のツールとなる

62

文献

- ▶ Campbell, G., Irony, T., Pennello, G., and Thompson, L. (2023). Bayesian statistics for medical devices: Progress since 2010. *Ther Innov Regul Sci* 57, 453-463.
- ▶ Diggle, P. J., Heagerty, P. J., Liang, K.-Y., and Zeger, S. L. (2002). *Analysis of Longitudinal Data*. Oxford: Oxford University Press.
- ▶ Dixon, D. O., and Simon, R. (1991). Bayesian subset analysis. *Biometrics* 47, 871-881.
- ▶ Dixon, D. O., and Simon, R. (1992). Bayesian subset analysis in a colorectal cancer clinical trial. *Stat Med* 11, 13-22.
- ▶ Higgins, J. P., and Spiegelhalter, D. J. (2002). Being sceptical about meta-analyses: A Bayesian perspective on magnesium trials in myocardial infarction. *Int J Epidemiol* 31, 96-104.
- ▶ Kandzari, D. E., Mauri, L., Koolen, J. J., Massaro, J. M., Doros, G., Garcia-Garcia, H. M., Bennett, J., Roguin, A., Gharib, E. G., Cutlip, D. E., Waksman, R., and Investigators, B. V. (2017). Ultrathin, bioresorbable polymer sirolimus-eluting stents versus thin, durable polymer everolimus-eluting stents in patients undergoing coronary revascularisation (BIOFLOW V): A randomised trial. *Lancet* 390, 1843-1852.

63

- ▶ Laird, N. M., and Ware, J. H. (1982). Random-effects models for longitudinal data. *Biometrics* 38, 963-974.
- ▶ Laptook, A. R., Shankaran, S., Tyson, J. E., Munoz, B., Bell, E. F., Goldberg, R. N., Parikh, N. A., Ambalavanan, N., Pedroza, C., Pappas, A., Das, A., Chaudhary, A. S., Ehrenkranz, R. A., Hensman, A. M., Van Meurs, K. P., Chalak, L. F., Khan, A. M., Hamrick, S. E. G., Sokol, G. M., Walsh, M. C., Poindexter, B. B., Faix, R. G., Watterberg, K. L., Frantz, I. D., 3rd, Guillet, R., Devaskar, U., Truog, W. E., Chock, V. Y., Wyckoff, M. H., McGowan, E. C., Carlton, D. P., Harmon, H. M., Brumbaugh, J. E., Cotten, C. M., Sanchez, P. J., Hibbs, A. M., Higgins, R. D., Eunice Kennedy Shriver National Institute of Child, H., and Human Development Neonatal Research, N. (2017). Effect of therapeutic hypothermia initiated after 6 hours of age on death or disability among newborns with hypoxic-ischemic encephalopathy: A randomized clinical trial. *JAMA* 318, 1550-1560.
- ▶ Mallinckrodt, C. H., Clark, W. S., and David, S. R. (2001). Accounting for dropout bias using mixed-effects models. *J Biopharm Stat* 11, 9-21.

64

- ▶ Matsuyama, Y., Sakamoto, J., and Ohashi, Y. (1998). A bayesian hierarchical survival model for the institutional effects in a multi-centre cancer clinical trial. *Stat Med* **17**, 1893-1908.
- ▶ Nakazato, H., Koike, A., Saji, S., Ogawa, N., and Sakamoto, J. (1994). Efficacy of immunochemotherapy as adjuvant treatment after curative resection of gastric cancer. Study group of immunochemotherapy with psk for gastric cancer. *Lancet* **343**, 1122-1126.
- ▶ Peto, R., Collins, R., and Gray, R. (1995). Large-scale randomized evidence: Large, simple trials and overviews of trials. *J Clin Epidemiol* **48**, 23-40.
- ▶ Quintana, M., Viele, K., and Lewis, R. J. (2017). Bayesian analysis: Using prior information to interpret the results of clinical trials. *JAMA* **318**, 1605-1606.
- ▶ Yusuf, S., and Flather, M. (1995). Magnesium in acute myocardial infarction. *BMJ* **310**, 751-752.
- ▶ Zhang, H., Lu, N., Feng, C., Thurston, S. W., Xia, Y., Zhu, L., and Tu, X. M. (2011). On fitting generalized linear mixed-effects models for binary responses using different statistical packages. *Stat Med* **30**, 2562-2572.
- ▶ 高井啓二, 星野崇宏, 野間久史. (2016). 欠測データの統計科学：医学と社会科学への応用. 東京: 岩波書店.

R example code

```
1 # 日本計量生物学会 2023年度計量生物セミナー
2 # 「階層ベイズモデルとその応用」
3 # 野間 久史 (統計数理研究所)
4 # 2023年12月14日
5
6 # R example code for performing Bayesian analyses of the meta-analysis examples in
7 # Higgins and Spiegelhalter (Int J Epidemiol. 31(1):96-104).
8 ###
9
10 install.packages("R2OpenBUGS") # Installation command
11 install.packages("coda")
12
13 ###
14
15 library("R2OpenBUGS")
16
17 # あらかじめ、以下のページから、OpenBUGSをインストールしておく必要があります。
18 # OpenBUGSは、デフォルトで、バックグラウンド稼働となりますので、すべての操作をR上で完結することができます。
19 # https://www.mrc-bsu.cam.ac.uk/software/bugs/openbugs/
20
21
22 ###
23
24 ## データセット(1): BMJのメタアナリシスにおける8試験のデータセット
25
26 rc <- c(2,23,7,1,8,9,3,118)
27 rm <- c(1,9,2,1,10,1,1,90)
28 nc <- c(36,135,200,46,148,56,23,1157)
29 nm <- c(40,135,200,48,150,59,25,1159)
30 N <- 8
31
32 madata8 <- list(rc=rc, rm=rm, nc=nc, nm=nm, N=N)
33
34 ###
35
36 ## データセット(2): ISIS-4 を除く14試験のデータセット
37
38 rc <- c(2,23,7,1,8,9,3,118,1,11,7,13,8,17)
39 rm <- c(1,9,2,1,10,1,1,90,0,6,1,5,4,4)
40 nc <- c(36,135,200,46,148,56,23,1157,21,75,27,33,122,108)
41 nm <- c(40,135,200,48,150,59,25,1159,22,76,27,23,130,107)
42 N <- 14
43
44 madata14 <- list(rc=rc, rm=rm, nc=nc, nm=nm, N=N)
45
46 ###
47
48 ## データセット(3): ISIS-4 を含む15試験のデータセット
49
50 rc <- c(2,23,7,1,8,9,3,118,1,11,7,13,8,17,2103)
51 rm <- c(1,9,2,1,10,1,1,90,0,6,1,5,4,4,2216)
52 nc <- c(36,135,200,46,148,56,23,1157,21,75,27,33,122,108,29039)
53 nm <- c(40,135,200,48,150,59,25,1159,22,76,27,23,130,107,29011)
54 N <- 15
55
56 madata15 <- list(rc=rc, rm=rm, nc=nc, nm=nm, N=N)
57
58 ###
59
60 ## BUGSのモデルコード(R上で、関数と定義して、R2OpenBUGSでそのまま使用することができます)
61
62 ## (a) Random-effects model; reference priorによる解析
63
64 rma01 <- function(){
65
66     for (i in 1:N) {
67         rc[i] ~ dbin(pc[i],nc[i]);
```

R example code

```
68     rm[i] ~ dbin(pm[i],nm[i]);
69     phi[i] <- logit(pc[i]);
70     logit(pm[i]) <- phi[i] + delta[i];
71     delta[i] ~ dnorm(mu, precision);           # Random effects
72     pc[i] ~ dunif(0,1);                       # Prior for pc
73 }
74
75 delta.new ~ dnorm(mu, precision);             # Predicted effect
76
77 mu ~ dnorm(0, 0.0001);                       # Reference prior for mu
78 tau ~ dunif(0, 100);
79
80 precision <- 1/(tau*tau);
81 tau.sq <- 1/precision;
82
83 OR <- exp(mu);
84
85 less0 <- min(delta.new, 0);                  # Calculate probabilities
86 prob0 <- 1 - equals(less0, 0);
87 less10 <- min(delta.new, -0.1054);
88 prob10 <- 1 - equals(less10, -0.1054);
89
90 }
91
92
93 ## (b) Random-effects model; sceptical priorによる解析
94
95 rma02 <- function(){
96
97     for (i in 1:N) {
98         rc[i] ~ dbin(pc[i],nc[i]);
99         rm[i] ~ dbin(pm[i],nm[i]);
100        phi[i] <- logit(pc[i]);
101        logit(pm[i]) <- phi[i] + delta[i];
102        delta[i] ~ dnorm(mu, precision);       # Random effects
103        pc[i] ~ dunif(0,1);                   # Prior for pc
104    }
105
106    delta.new ~ dnorm(mu, precision);          # Predicted effect
107
108    mu ~ dnorm(0, 32.69);                     # Sceptical prior for mu
109    tau ~ dunif(0, 100);
110
111    precision <- 1/(tau*tau);
112    tau.sq <- 1/precision;
113
114    OR <- exp(mu);
115
116    less0 <- min(delta.new, 0);               # Calculate probabilities
117    prob0 <- 1 - equals(less0, 0);
118    less10 <- min(delta.new, -0.1054);
119    prob10 <- 1 - equals(less10, -0.1054);
120
121 }
122
123
124 ###
125
126 ## 初期値の設定
127
128 mainits <- function(){
129     list(mu = 0, tau = 1)
130 }
131
132
133 ## OpenBUGSによるMCMC: (1) Reference priorによる解析: 14 trials
134
135 rmaout.01 <- bugs(data=madata14,
```

```
136     inits=mainits,
137     parameters.to.save=c("OR", "mu", "tau", "delta.new", "prob0", "prob10"),
138     model.file=rma01,
139     n.chains=1,
140     n.iter=22000,
141     n.burnin=2000)
142
143
144 ## OpenBUGSによるMCMC: (2) Reference priorによる解析: 14 trials + ISIS-4
145
146 rmaout.02 <- bugs(data=madata15,
147                 inits=mainits,
148                 parameters.to.save=c("OR", "mu", "tau", "delta.new", "prob0", "prob10"),
149                 model.file=rma01,
150                 n.chains=1,
151                 n.iter=22000,
152                 n.burnin=2000)
153
154
155 ## OpenBUGSによるMCMC: (3) Sceptical priorによる解析: 14 trials
156
157 rmaout.03 <- bugs(data=madata14,
158                 inits=mainits,
159                 parameters.to.save=c("OR", "mu", "tau", "delta.new", "prob0", "prob10"),
160                 model.file=rma02,
161                 n.chains=1,
162                 n.iter=22000,
163                 n.burnin=2000)
164
165
166 ## OpenBUGSによるMCMC: (4) Sceptical priorによる解析: 14 trials + ISIS-4
167
168 rmaout.04 <- bugs(data=madata15,
169                 inits=mainits,
170                 parameters.to.save=c("OR", "mu", "tau", "delta.new", "prob0", "prob10"),
171                 model.file=rma02,
172                 n.chains=1,
173                 n.iter=22000,
174                 n.burnin=2000)
175
176
177
178 ## 結果のサマリーの出力
179
180 rmaout.01$summary
181 rmaout.02$summary
182 rmaout.03$summary
183 rmaout.04$summary
184
185
186
187 ###
188
189 ## coda packageのツールによる収束診断など
190
191 library("coda")
192
193 crmaout.03 <- bugs(data=madata14,
194                  inits=mainits,
195                  parameters.to.save=c("mu", "tau", "delta.new"),
196                  model.file=rma02,
197                  n.chains=5,
198                  n.iter=5000,
199                  n.burnin=2000,
200                  codaPkg=TRUE) # "codaPkg=TRUE"
                                の引数を入れておくと、codaとの連携ができるオブジェクトが出力されます。
201
202 ermaout.03 <- read.bugs(crmaout.03)
```

R example code

```
203
204 summary(ermaout.03)
205 HPDinterval(ermaout.03) # 最高事後密度区間
206
207 batchSE(ermaout.03) # バッチ標準誤差
208 autocorr.diag(ermaout.03) # 自己相関係数による診断
209
210 plot(ermaout.03) # Trace plot and density estimation
211
212 gelman.diag(ermaout.03) # Gelman-Rubin診断
213 gelman.plot(ermaout.03)
214
215 geweke.diag(ermaout.03) # Gewekeの収束診断
216 geweke.plot(ermaout.03)
217
218
```