SIS 2017
Statistics and Data Science: new challenges, new generations

28–30 June 2017
Florence (Italy)

Proceedings of the Conference of the Italian Statistical Society

edited by
Alessandra Petrucci
Rosanna Verde

FIRENZE UNIVERSITY PRESS
2017

SIS 2017. Statistics and Data Science: new challenges, new generations : 28-30 June 2017 Florence (Italy) : proceedings of the Conference of the Italian Statistical Society / edited by Alessandra Petrucci, Rosanna Verde. – Firenze : Firenze University Press, 2017.
(Proceedings e report ; 114)

http://digital.casalini.it/9788864535210

ISBN 978-88-6453-521-0 (online)

Peer Review Process
All publications are submitted to an external refereeing process under the responsibility of the FUP Editorial Board and the Scientific Committees of the individual series. The works published in the FUP catalogue are evaluated and approved by the Editorial Board of the publishing house. For a more detailed description of the refereeing process we refer to the official documents published on the website and in the online catalogue of the FUP (www.fupress.com).

Università degli Studi di Firenze
Firenze University Press
via Cittadella, 7, 50144 Firenze, Italy
www.fupress.com

Causal analysis of Cell Transformation Assays

Analisi causale dei Cell Transformation Assay

Federico Mattia Stefanini

Abstract A Cell Transformation Assay (CTA) is an *in vitro* method to test a chemical for carcinogenicity. In a recent contribution from an international expert group created to improve the analysis of BALB/c 3T3 CTA data, two classes of models in the frequentist paradigm were recommended. Here a Bayesian model for potential outcomes is developed to estimate the causal effect of some concentrations of a candidate carcinogen on counts of *foci* growing within Petri dishes. The reanalysis of an actual case study is performed to illustrate some limitations of current models and the main features of the proposed approach.

Abstract *(in Italian)* Il Cell Transformation Assay (saggio di trasformazione cellulare) è un metodo *in vitro* per saggiare la carcinogenicità di una sostanza chimica. In un recente contributo di un gruppo di esperti internazionali, creato per migliorare l'analisi dei dati provenienti dal saggio con la linea cellulare BALB/c 3T3, sono stati raccomandati due classi di modelli nel paradigma frequentista. In questo lavoro viene sviluppato un modello Bayesiano per i risultati potenziali allo scopo di stimare l'effetto causale di alcune concentrazioni di un candidato cancerogeno sul conteggio dei *foci* che crescono entro capsule Petri. La rianalisi di un caso di studio reale viene realizzata per illustrare alcune limitazioni dei metodi correnti e le principali caratteristiche del modello proposto.

Key words: Bayesian causal model, potential outcomes

1 Introduction

It has been estimated that annual cancer incidence will rise from 14 million in year 2012 to 22 within the next 2 decades [1]. Chemical carcinogenicity is defined as

Department of Statistics, Computer Science, Applications
University of Florence
e-mail: stefanini@disia.unifi.it

Alessandra Petrucci, Rosanna Verde (edited by), *SIS 2017. Statistics and Data Science: new challenges, new generations. 28-30 June 2017 Florence (Italy). Proceedings of the Conference of the Italian Statistical Society*
ISBN (online) 978-88-6453-521-0 (online), CC BY 4.0, 2017 Firenze University Press

the ability of a chemical substance, or a mixture of chemical substances, to induce cancer or to increase its incidence. Given the role played by environmental and chemical exposures, no wonder that the evaluation of chemical carcinogenicity has become a leading task in public health risk assessment during the last decades.

Cell Transformation Assays (CTAs) are a family of *in vitro* methods for the identification of potential chemical carcinogens. It has been shown that CTAs nicely correlate with rodent bioassay, which is considered the standard approach for carcinogenicity testing [2]. The endpoint assessed in CTAs is the progression of cultured cells from immortality to tumorigenicity, as evidenced by formation of *foci* of multilayered and disorganised cells, growing over the surrounding regular cell monolayer. Therefore, the number of fully transformed cell colonies, called type III *foci*, grown within a Petri dish (experimental unit) after 4 weeks from treatment with the chemical under testing is the outcome of primary interest.

In the following, a Bayesian model for potential outcomes is developed to estimate the causal effect of different concentrations of a chemical in a CTA.

2 A Bayesian model

The case study here considered is a CTA experiment performed to test o-toluidine (CAS chemical registry number # 636-21-5). A total of eigth different concentrations and the negative control were considered, and they are (μg/ml): 0 (negative control, $i = 0$), 20 ($i = 1$), 100 ($i = 2$), 200 ($i = 3$), 500 ($i = 4$), 800 ($i = 5$), 1000 ($i = 6$), 1200 ($i = 7$), 1750 ($i = 8$). A total of 90 Petri dishes containing BALB/c 3T3 cells sampled at the same passage from the original cell culture were treated after random assignment of each concentration to $n_i = 10$ dishes (replicates) for each i. All experimental units received protocol ingredients taken from the same batch, including medium and serum. After 4 weeks from treatment, Petri dishes were visually scored under a light microscope and the number of type III *foci* within each dish counted.

Following Rubin's framework for causal inference [3], potential outcomes are introduced for every treatment and experimental unit under consideration. Let $Y_k^{<i>}$ be random variables representing the potential number of *foci* within Petri dish $k = 1, \ldots, n$ under treatment (concentration) $i = 0, 1, \ldots, L$. A plausible size for the sample space $\Omega_{Y<i>}$ is around 30, thus $\Omega_{Y<i>} = \{0, 1, \ldots, 30\}$, because the available physical space on a Petri dish is limited.

The potential outcomes referred to concentration i define the vector $Y^{<i>}$. Let $W = (W_1, \ldots, W_n)^T$ be the vector of indicators of treatment assignment, with sample space $\Omega_{W_i} = \{0, \ldots, L\}$. Given that CTAs belong to the class of randomized experiments, the assignment mechanism is ignorable and characterized by unit-level probability of treatment assignment in the interval $(0, 1)$, in particular the probability mass function of vector W that represents the assignment mechanism is:

$$p(W \mid Y^{<0>}, \ldots, Y^{<L>}) = \binom{n}{n_1 n_2 \ldots n_L}^{-1} \tag{1}$$

for all W satisfying $\sum_{k=1}^{n} I_i(W_k) = n_i$ for each i. Under row (unit) exchangeability of matrix $(Y^{<0>}, \ldots, Y^{<L>})$ the joint distribution of potential outcomes is:

$$p\left(Y^{<0>}, \ldots, Y^{<L>}\right) = \int \prod_{k=1}^{n} p(Y_k^{<0>}, \ldots, Y_k^{<L>} \mid \theta) p(\theta)\, d\theta \tag{2}$$

where θ is a vector of model parameters belonging to the parameter space Θ.

The elicitation of conditional distributions for potential outcomes given model parameters (eq. 2), the so called science, should take into account the main processes driving the emergence of *foci*. Even if it is not carcinogenic, a chemical may exert a toxic effect on cultured cells, thus causing a reduction in the final number of type III *foci*. If a chemical is carcinogenic then it is expected to stimulate the emergence of *foci*, but this driving force also depends on concentration: too low doses are ineffective, too high doses are often cytotoxic. Despite that concentrations are selected to be within a convenient range, it is quite difficult to anticipate any correlation between potential outcomes. For these reasons, conditional independence among potential outcomes is here assumed:

$$p\left(Y_1^{<0>}, \ldots, Y_k^{<L>}, \theta\right) = \prod_{i=0}^{L} p(Y_k^{<i>} \mid \theta_i) p(\theta_i) \tag{3}$$

where the joint distribution of model parameters is factorized into marginally independent subvectors, $\theta = (\theta_0, \ldots, \theta_i, \ldots, \theta_L)$.

At the end of the experiment, the vector of observed potential outcomes is:

$$Y^{<obs>} = \left\{ \left(\ldots, \sum_{i=0}^{L} Y_k^{<i>} I_i(W_k), \ldots \right)^T : \; k = 1, \ldots, n \right\}$$

while the collection of vectors $C^{<mis>} = \{Y^{<mis,0>}, \ldots, Y^{<mis,L>}\}$ with

$$Y^{<mis,i>} = \left\{ \left(\ldots, Y_k^{<i>}, \ldots \right)^T : \; W_k \neq i, \;\; k = 1, \ldots, n, \right\}$$

and $i = 0, \ldots, L$, contains missing potential outcomes. The conditional predictive distribution $p(C^{<mis>} \mid Y^{<obs>}, W)$ is exploited to impute missing values.

Three causal estimands of particular interest are finite sample averages of indicators, with $i \geq 1$: the probability of positive effect on the treated (PPET) units,

$$\tau_{PPET}^{<i>} = \sum_{k=1}^{n} I_{\{y^{<i>} - y^{<0>} > 0 \,\wedge\, W_k = i\}} (Y_k^{<i>}, Y_k^{<0>}, W_k) / \sum_{k=1}^{n} I_i(W_k),$$

the probability of null effect on the treated (PNET) units, where the indicator of the event is $I_{\{y^{<i>}-y^{<0>}=0 \wedge W_k=i\}}(Y_k^{<i>}, Y_k^{<0>}, W_k)$, the probability of cytotoxic effect on the treated (PCET) units, where $I_{\{y^{<i>}-y^{<0>}<0 \wedge W_k=i\}}(Y_k^{<i>}, Y_k^{<0>}, W_k)$.

Further structure may be imposed on the model after borrowing context and assumptions settled by an international expert group (European Centre for the Validation of Alternative Methods, ECVAM), in particular:

1. the number of Petri dishes at each concentration is typically 10 (never below 9);
2. the number of levels for the concentration typically ranges from 3 to 7;
3. *focus*-inducing chemicals are expected to show non-monotone dose-concentration relationships, mostly due to cytotoxicity at higher concentrations;
4. positive controls are not informative;
5. at small concentrations the empirical distribution of counts may be degenerate, typically at zero;
6. concentrations have to be considered as levels of a qualitative factor, although originally on a quantitative scale (μg/ml).

ECVAM's experts recommended two tentative classes of models, the first one is a Normal model for Nishiyama-transformed counts, $x = \sqrt{y} + \sqrt{1+y}$, and the second one is a Negative Binomial model for original counts. Unfortunately, the two recommended classes did not seem suited to our case study (Section 3).

ECVAM's committee proposed two family of distributions allowing asymmetry (original scale) and smooth changes of probability value in contiguous (transformed) counts. By introducing latent variables $X_k^{<i>} \sim Beta(x \mid \alpha_i, \beta_i)$ in the Beta family, we essentially maintained the original belief while gaining in flexibility: values of variance smaller than the mean became possible. The probability of observing count $y_{i,j}$ is thus $P[Y_j^{<i>} = y] = \int_{y/31}^{(y+1)/31} Beta(x \mid \alpha_i, \beta_i)dx$. A weakly informative initial distribution was elicited for marginally independent model parameters, with $\alpha_i \sim Uniform(1, 1000)$ and $\beta_i \sim Esponential(0.01)$, $i = \{0, \ldots, L\}$.

3 Results and discussion

Computations were performed in R[1] using RStudio[2] and the following packages[3]: *MASS, fitdistrplus, rjags,coda,knitr*.

In the first step of the analysis, Normal models for Nishyama-transformed counts were considered. Note that after transformation, null counts are mapped to 1. From unbiased point estimates of model parameters at each concentration i, plug-in estimates of probability values $\hat{P}[X_i < 1 \mid \hat{\mu}_i, \hat{\sigma}_i^2]$ were calculated, and for concentrations from 0 to 200 they resulted well above 0.15, that is: 0.1904, 0.2265, 0.2673, 0.2734.

[1] `https://www.R-project.org/`

[2] `http://www.rstudio.com`

[3] `https://cran.r-project.org/web/packages/`

For concentrations equal to 500 and 800 estimates were 0.1470 and 0.0550 respectively. Only for the last 3 concentrations the point estimates were below 0.02. Fitting distributions by maximum likelihood always reduced the the above estimated probability values of about 0.01. Quantile-quantile plots (not shown), even if based on just 10 observations, detected clear departures of Nishiyama-transformed counts from normality for all concentrations smaller than 800.

Note that no concentration showed counts all equal to zero (or to one), an event of appreciable probability in many case studies, therefore it was possible to obtain the unbiased point estimate of the variance. For this reason, we did not consider the recommended artificial increase of sample size by one observation equal 1, an action that would have determined an increase of sample size of about 10% at each concentration. Furthermore, in case all observations are equal to one, such artificial change of observed counts is not even uniquely defined. Given the role played by the predictive distribution in the Bayesian causal model, and therefore by the model for observations, the Normal model for Nishiyama-transformed counts seemed unsatisfactory and therefore was not considered further.

In the second step, we considered the class of Negative Binomial models. The optimization of likelihood functions at each concentration often failed due to divergence of the scale parameter towards infinity. Even upon termination, it was sometimes impossible to calculate standard errors, or in other cases estimated values were huge. Similar failures were observed using other algorithms, for example iterated moment matching. Indeed, a small sample size at each concentration ($n_i = 10\ \forall i$) and the presence of sample variances often smaller than sample averages made the optimization hard, given that the Negative Binomial family is not suited to underdispersed count data. All things considered, the class of Negative Binomial models seemed unsatisfactory for the case study at hand and therefore it was not considered further.

The proposed Beta latent model was fitted by MCMC: a sample of 1×10^5 realizations from the final distribution of model parameters was collected after thinning one chain by 4. The initial burn-in consisted of $10'000$ iterations. Values of difference between pair of counterfactuals in the numerator of $\tau_{PPET}^{<i>}$, $i = 1, \ldots, 8$ were saved and further processed to obtain their distribution conditioned to observed outcomes at each concentration. One-chain output diagnostics did not suggest lack of convergence. In Table (1), estimated probability of carcinogenic/null/cytotoxic effects are shown, based on a sample of 1×10^6 imputed counterfactuals for each concentration. The odds of a carcinogenic effect of o-toluene at concentrations 1000 and 1200 is about ten. The 5% quantile of the distribution of odds for a carcinogenic effect of o-toluene is equal to 2.3333 at both concentrations 1000 and 1250, thus they are both well above 1.

Table 1 Estimated probability values of causal effects.

Concentration $< i >$	20	100	200	500	800	1000	1200	1750
$\tau_{PCET}^{<i>}$	0.403	0.454	0.441	0.324	0.194	0.032	0.024	0.107
$\tau_{PNET}^{<i>}$	0.386	0.417	0.391	0.365	0.254	0.067	0.053	0.147
$\tau_{PPET}^{<i>}$	0.211	0.129	0.168	0.311	0.552	0.901	0.923	0.746
odds $\frac{\tau_{PPET}^{<i>}}{(1-\tau_{PPET}^{<i>})}$	0.266	0.148	0.202	0.451	1.232	9.101	11.987	2.937

4 Conclusions

We developed a Bayesian causal model based on latent Beta distributions to overcome limitations found in alternative proposals when applied to the o-toluene case study. Instead of exploiting beliefs by adding virtual observations, prior information was formally introduced. Causal estimands like PPET were formulated to restrain the assessment by excluding the magnitude of differences, but motivations of this choice will be detailed elsewhere. Further similar estimands might address effects due to the increase of concentration, an important issue for applications not reported in this work.

The limited sample size and the small number of concentrations characterizing a typical CTA make the assessment of model assumptions hard. It is clear that alternative, and more general, latent models might be formulated, for example one that allows for rare but very extreme counts. These outliers were absent in our case study but they are not so rarely observed in CTAs. Replicates of the same experiment performed on a chemical in different laboratories represent an opportunity to increase sample size, at least after properly considering transportability.

Finally, a battery of positive controls made by known genotoxic and nongenotoxic carcinogens, for example 3-methylcolantrene, could be introduced to study the variability in the response of experimental units to treatments.

Acknowledgements The author thanks Chiara Urani e Giulia Callegaro for their valuable expertise and many fruitful discussions on CTAs. This work is partially supported by University of Florence, frame 'Disegno e analisi di studi sperimentali e osservazionali per le decisioni'.

References

1. Stewart, B. W., C. P. Wild: World Cancer Report 2014. OCLC: 908606220. Lyon: International Agency for Research on Cancer/World Health Organization (2014).
2. Creton, S., Aardema, M.J., Carmichael, P.L., Harvey, J.S., Martin, F.L., Newbold, R.F., O'Donovan, M.R., Pant, K., Poth, A., Sakai, A., Sasaki, K., Scott, A.D., Schechtman, L.M., Shen, R.R., Tanaka, N., Yasaei, H.: Cell transformation assays for prediction of carcinogenic potential: state of the science and future research needs. Mutagenesis **27**, 93–101 (2012).
3. Rubin, D.: Bayesian Inference for Causal Effects: The Role of Randomization. The Annals of Statistics **6**, 34–58, (1978).