ISSN: 2317-0840

The *Salmonella*/microsome mutagenicity index revisited: statistical performance evaluation and modified measures

Davi Butturi-Gomes[1†], Silvio S. Zocchi[2]

[1] *Departamento de Matemática e Estatística, Campus Santo Antônio, Universidade Federal de São João Del-Rei (DEMAT/UFSJ).*
[2] *Programa de Pós-graduação em Estatística e Experimentação Agronômica, Departamento de Ciências Exatas, Escola Superior de Agricultura "Luiz de Queiroz", Universidade de São Paulo (ESALQ/USP). E-mail: sszocchi@usp.br.*

Abstract: *The Ames test is a widely used, short-term mutagenicity assay for environmental and human health risk assessment and its results are usually expressed by means of a simplistic measure, namely the mutagenicity index (MI). In this study, we aimed at evaluating its statistical properties, its advantages and disadvantages (point and interval estimation performances) in comparison to three new, similar measures (MI_m, $\tilde{MI}$ and $\tilde{MI}_m$). We used computer simulations of realistic, protocol-recommended scenarios, and illustrated our methods with an observed, real dataset. Our major findings are the following: (i) all four measures are generally unbiased; (ii) both bias and variance raises with overdispersion; and (iii) only the MI_m is guaranteed to meet the nominal confidence level in interval estimation. Along with other recommendations, we present a discussion on the implications of the misuse of any these indices.*

Keywords: Ames test; Bias; Bootstrap; Confidence intervals; Overdispersion

Resumo: *O teste de Ames é um ensaio biológico amplamente utilizado e de curto prazo para avaliação de riscos ambientais e de saúde, cujos resultados são geralmente expressado por meio de uma medida simplista, s saber, o índice de mutagenicidade (MI). Neste estudo, nosso objetivo foi avaliar as propriedades estatísticas desse índice, suas vantagens e desvantagens (desempenhos de estimação pontual e intervalar) em comparação a outras três novas medidas similares (MI_m, $\tilde{MI}$ e $\tilde{MI}_m$). Utilizamos simulações em computador de cenários realistas seguindo os protocolos de aplicação comumente adotados, e ilustramos nossos métodos com um conjunto de dados reais. Nossos principais resultados foram: (i) todas os quadro índices são em geral não-viesados; (ii) tanto o viés como a variância aumenta com a superdispersão; e (iii) apenas MI_m garantiu o alcance do nível nominal de confiança na estimação intervalar. Em conjunto a outras recomendações, apresentamos uma discussão de possíveis consequências em situações em que esses índices forem mal utilizados.*

Palavras-chave: Teste de Ames; Viés; Bootstrap; Intervalo de confiança; Superdispersão

[†]Correspondent author: davibg@ufsj.edu.br.

Introduction

The Salmonella/microsome assay, also known as Ames Test (AT) (Ames et al. 1973a; Ames et al. 1973b), is a widely accepted experimental procedure in the scientific community, due to its relatively low cost and to its simplicity and readiness of results. Although several statistical models are available for AT data (see Kim and Margolin (1999) for a comprehensive review on this issue), the mutagenicity index (MI) is perhaps the most used measure whenever biologists are in charge of performing the assay, analyzing its results and, sometimes, even compare the mutagenic effects of different substances.

The formula for computing the MI is

$$\mathrm{MI}_k = \frac{\frac{1}{n_k}\sum_{j=1}^{n_k} y_{kj}}{\frac{1}{n_0}\sum_{i=1}^{n_0} y_{0i}} = \frac{\bar{y}_k}{\bar{y}_0}, \tag{1}$$

where n_k is the number of replicates used for the k-th dose, n_0 is the number of replicates used for the negative control, y_{kj} is the number of revertants observed in the j-th replicate of the k-th dose and y_{0i} is the number of revertants observed in the i-th replicate of the negative control.

Despite its popularity, using the MI in Equation (1) has a few downsides researchers often neglect. Firstly, we must assume that the substance of interest is non-toxic in the chosen doses (or concentrations) for the assay in order to successfully use the MI. Furthermore, a common phenomenon that arises whenever dealing with count data is known as overdispersion (Edler 1992; Hinde and Demétrio 1998; Molla and Muniswamy 2012), which, when unaddressed, may further compromise both point and interval estimators of the MI.

More importantly, perhaps, is that the MI statistical behavior is questionable due to the small number of experimental units suggested by AT protocols (Mortelmans and Zeiger 2000; OECD 1997; Umbuzeiro et al. 2009). Since the early works of Katz (1978, 1979), who proposed test statistics and minimal sample sizes using an asymptotic theoretical approach, the scientific community is already aware that effectively detecting the 2- or 3-fold rules for mutagenesis in Salmonella is mostly impractical (see Cariello and Piegorsch (1996) for exact error rates of such measures).

Said statistical flaws and misuses of the MI are closely related to type-I and type-II errors (respectively, the probability of rejecting a true hypothesis and the probability of not rejecting a false one) and may lead to very serious, practical implications, such as human and environmental risks. Especially, these may be the cases encountered where the observed MIs point to either weakly positive or negative mutagenic effects, according to the researchers' fold rule criterion (see, for instance, the works of Akhtar et al. (2016), Barbee et al. (1996), Gupta et al. (2014), López et al. (1999), Romero et al. (1997) and of Takemura et al. (2010)).

Thus, given the importance of the MI for environmental and human health risk assessment, its constant misuses and eventual simply plain wrong interpretations, we aimed at deriving similar, alternative measures to the MI, and at evaluating the performances of point and interval estimation of the usual MI and of these modified measures, under realistic scenarios of overdispersion and sampling effort. We also provided a short data analysis to illustrate our results.

Methods

In order to evaluate the statistical performance of the MI, we performed computer simulations using Monte Carlo methods, under realistic, protocol-recommended laboratory scenarios. We carried out all simulations in the R software (R CORE TEAM, 2015). The details of our procedures are presented below.

Since we suspected that the MI may be a problematic measure, we derived a modification, namely the modified mutagenicity index (MI_m), whose formula for the k-th dose used in an assay is given by

$$\mathrm{MI}_{mk} = \frac{1}{n_0 n_k} \sum_{i=1}^{n_0} \sum_{j=1}^{n_k} \frac{y_{kj}}{y_{0i}} \tag{2}$$

where n_0, n_k, y_{0i} and y_{kj} are the same quantities as those described for Equation (1). The index in Equation (2) emerges naturally whenever dealing with ratios of unpaired observations and, thus, we included the MI_m in our simulations along with the usual MI. One often computes a confidence interval (CI) for an observed MI value using the normal-theory approach.

Nevertheless, due to the count nature of the response variable (or proportions of such counts) and to the rather small sample sizes, said asymptotic approximations are hardly met in practice. Thus, in order to overcome these issues with simple normal CIs, we applied nonparametric bootstrap and jackknife methods for bias correction and both, normal and percentile-based, bootstrap confidence intervals (Efron and Tibshirani 1993).

Since we are not making any assumptions regarding the underlying distributions of neither MI nor MI_m, it may be wise to consider median-based measures as well. Thus, we also included the following mutagenicity indices

$$\tilde{\mathrm{MI}}_k = \frac{Q_2(y_{k1}, y_{k2}, \ldots, y_{kn_k})}{Q_2(y_{01}, y_{02}, \ldots, y_{0n_0}))} = \frac{\tilde{y}_k}{\tilde{y}_0} \tag{3}$$

and

$$\tilde{\mathrm{MI}}_{mk} = Q_2 \left(\frac{y_{k1}}{y_{01}}, \frac{y_{k2}}{y_{01}}, \ldots, \frac{y_{kn_k}}{y_{01}}, \ldots, \frac{y_{k1}}{y_{0n_0}}, \frac{y_{k2}}{y_{0n_0}}, \ldots, \frac{y_{kn_k}}{y_{0n_0}} \right) \tag{4}$$

where $Q_2(\cdot)$ is the second quantile of the sample (i.e., the sample median) and the other quantities have already been defined. A common dispersion measure for the median is known as dispersion coefficient (DC) and, although useful for data visualization, one usually does not use DCs to obtain CIs for the median. Instead, as in p. 120 of Zar (2010), we used binomial CIs, both exact and asymptotic. Finally, we applied nonparametric bootstrap and jackknife methods for bias correction and percentile-based CI estimation for $\tilde{\mathrm{MI}}$ and $\tilde{\mathrm{MI}}_m$ (Equations (3) and (4), respectively).

For the simulations, we used one of the most recommended experimental scheme: three replicates of each dose (of a hypothetical substance) plus the negative control, in a completely random design (CRD) (Umbuzeiro et al. 2009). For the response variable (number of revertants in each experimental unit), we supposed both Poisson and negative binomial (NB-2, see Lawless (1987), for instance) distributions, to accommodate overdispersion, since it is a common phenomenon that arises with count data and particularly with the AT (Bernstein et al. 1982; Breslow 1984). We considered four levels of overdispersion (in some arbitrary unit): 0 (Poisson distribution), 1 (NB-2 with size parameter $k = 100$), 2 (NB-2 and $k = 50$) and 4 (NB-2 and $k = 25$).

In all scenarios, we varied the "true MI"(herein denoted by mutagenic effect or simply δ) ranging from 1 to 9 in steps of 0.08, totaling 101 values for δ per simulation. For the negative control, we considered the population expectation to be 110 colonies, which reflects the expected number of spontaneous revertants for the TA100 strain (Mortelmans and Zeiger 2000). We performed 10,000 simulations for each level of dispersion, thus totaling 40,000 simulations. Within each simulation, we used 1,000 bootstrap samples.

Results and Discussion

We found a clear tendency for the bias of all point estimators to increase with overdispersion level, though they appeared robust against δ (Figure 1). Nevertheless, the bias of any estimator is rather small, up to an overestimation of 5% its true value, particularly for the MI_m (Equation (2)) in the scenarios simulated using NB-2 with $k = 25$. The resampling estimation methods generally did not produce neither increase nor decrease of the percentage bias, except maybe for the jackknifed $\tilde{\mathrm{MI}}_m$, whose biases were robust against overdispersion.

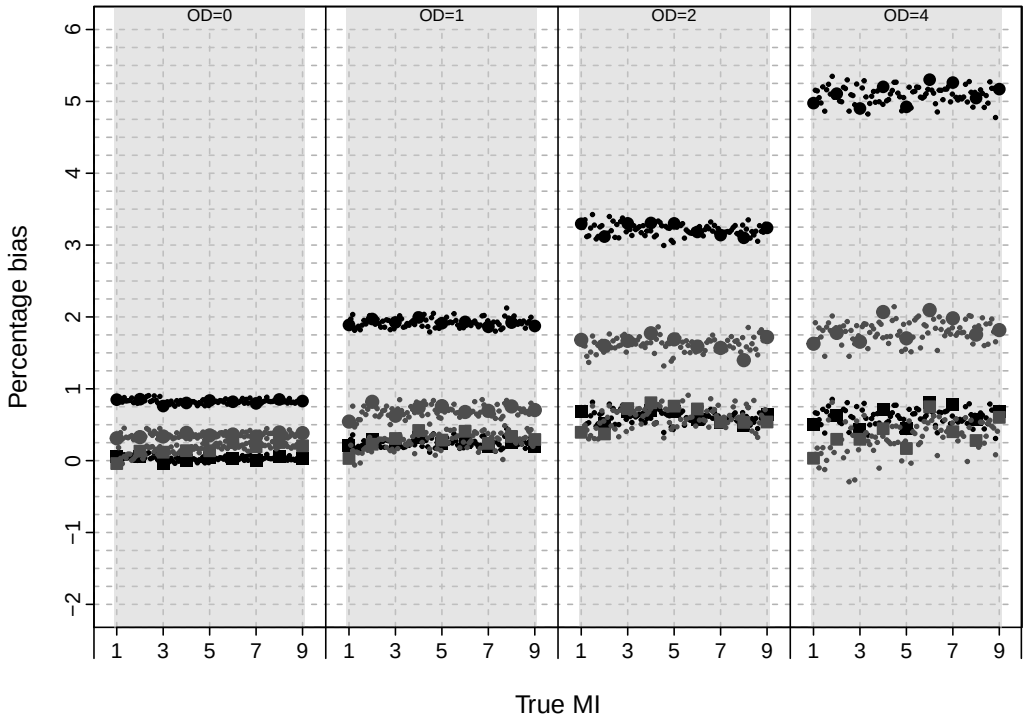

Figure 1: Percentage bias of different estimators (averaged by estimation method) according to the true MI. Symbols: ●MI; ■MI_m; ●$\tilde{\mathrm{MI}}$; ■$\tilde{\mathrm{MI}}_m$. Overdispersion levels: OD=0 (Poisson), OD=1 (negative binomial with $k = 100$), OD=2 (negative binomial with $k = 50$) and OD=4 (negative binomial with $k = 25$).

Regarding the mean squared error (MSE), the best performances (smaller values of MSE) were for the MI and MI_m, despite the fact that the MSE increases with both overdispersion level and δ. Although not by much, the jackknife version of all indices showed the worst point estimation performances using the MSE criterion (Figure 2).

The precision of all evaluated CIs decrease with both, mutagenic effect and overdispersion. Percentile-based and normal bootstrapped CIs were the most precise and most robust against the choice of mutagenic index, while asymptotic normal, exact binomial and asymptotic binomial varied most in precision (Figure 3).

The accuracy of all CIs is relatively robust against overdispersion and mutagenic effect, although only asymptotic normal and jackknifed normal CIs for the MI_m met the minimum, nominal confidence level of 95% (Figure 4).

The analysis quinoline data (Margolin et al. 1981) is presented in Table 1 (interval estimation). If one would invert all CIs to test a hypothesis such as $\mathrm{H}_0 : \mathrm{MI} \leq 2$ vs. $\mathrm{H}_1 : \mathrm{MI} > 2$ (i.e., the 2-fold rule hypothesis test), then they would find no evidence of mutagenic action of quinoline over TA98 strain, for all evaluated doses and all mutagenicity indices, at some $0.01 < \alpha < 0.05$ nominal significance level and an appropriate multicomparison correction.

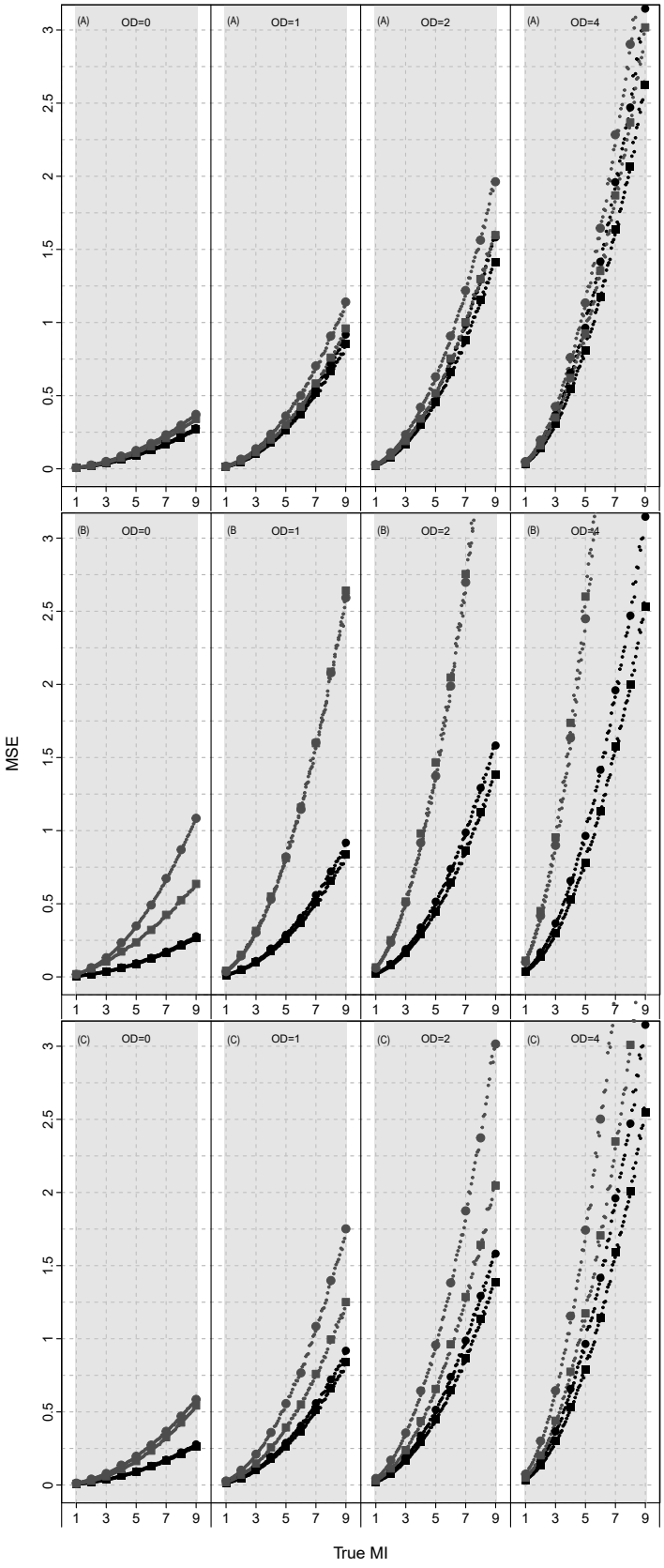

Figure 2: Mean squared error according to the true MI. Symbols: ●MI; ■MI_m; ●$\tilde{\text{MI}}$; ■$\tilde{\text{MI}}_m$. Estimation methods: (A) standard (plug-in) estimator; (B) bias-corrected jackknife; (C) bias-corrected bootstrap. Overdispersion levels: OD=0 (Poisson), OD=1 (negative binomial with $k = 100$), OD=2 (negative binomial with $k = 50$) and OD=4 (negative binomial with $k = 25$).

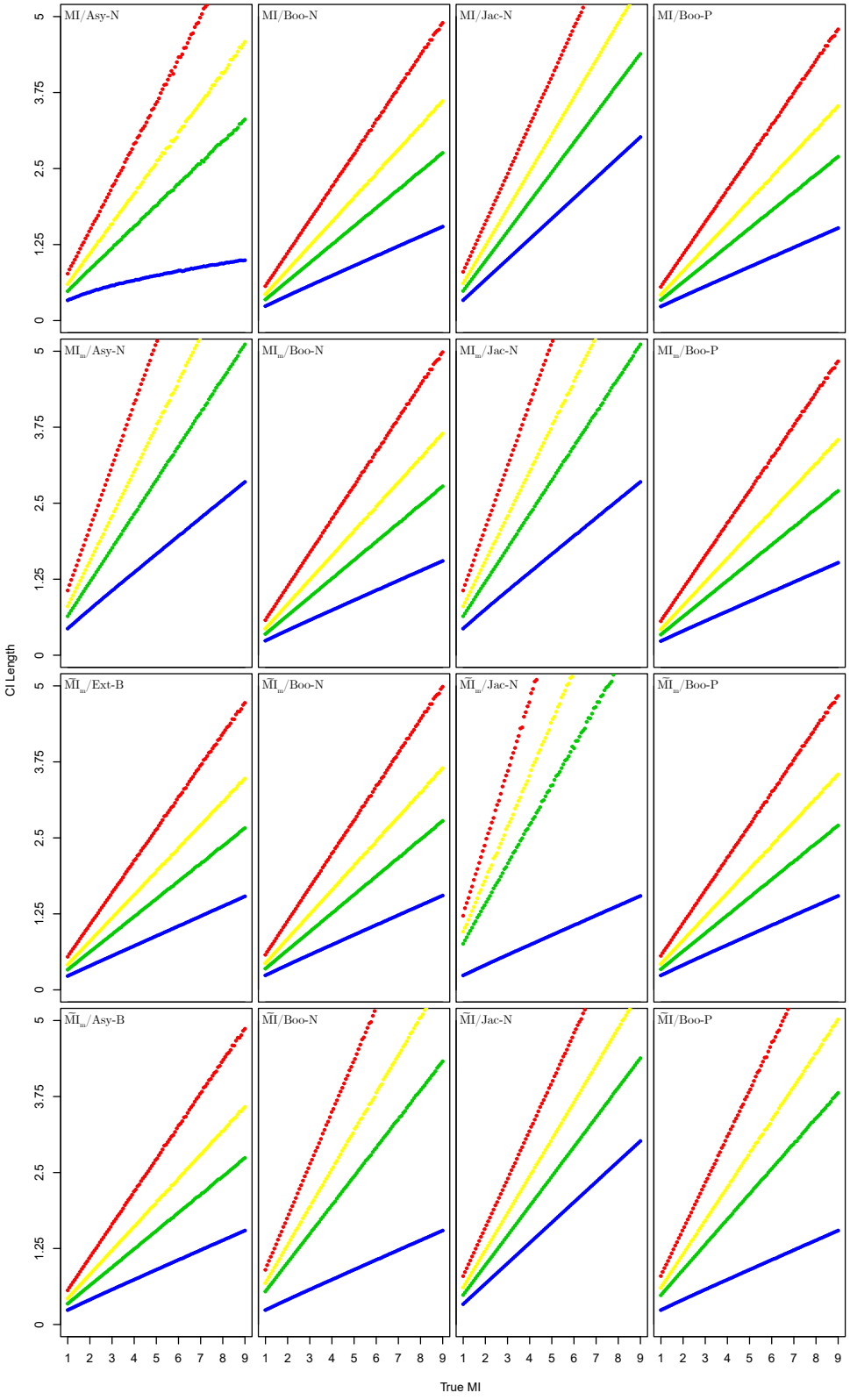

Figure 3: CI length (precision) of different interval estimators. Asy-N: asymptotic normal; Ext-B: exact binomial; Asy-B: asymptotic binomial; Jac-N: jackknifed normal; Boo-N: bootstrapped normal; Boo-P: percentile-based bootstrap interval. Blue: OD=0 (Poisson); green: OD=1 (negative binomial with $k = 100$); yellow: OD=2 (negative binomial with $k = 50$); and red: OD=4 (negative binomial with $k = 25$).

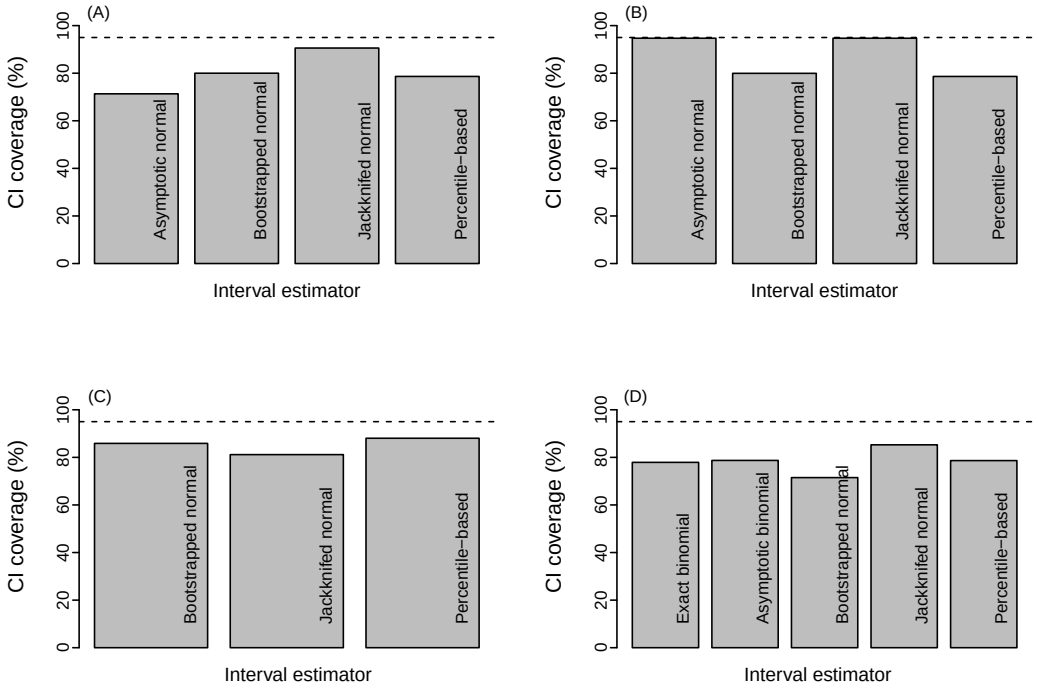

Figure 4: Effective level of confidence (accuracy) of different interval estimators averaged by mutagenicity effect. Each panel contains the available methods of interval estimation for (A) MI; (B) MI_m; (C) $\tilde{\mathrm{MI}}$; and (D) $\tilde{\mathrm{MI}}_m$.

The results corroborate our initial hypotheses (i.e., biases and variances increasing with overdispersion and typical lower precision in scenarios of higher mutagenic effect) and some of them are useful in choosing a particular mutagenicity index and its estimation method.

On the one hand, if one is interested in point estimation, jackknifing the $\tilde{\mathrm{MI}}_m$ is a good choice if overdispersion information in the dataset is either unavailable or cannot be extracted. On the other hand, if one finds it important to minimize the variance, allowing for somewhat more biased point estimation, then they should certainly use the bootstrapped MI_m.

If one is interested in interval estimation, using the asymptotic-normal interval for the MI_m is the only reasonable choice, since it meets the nominal confidence level, did not show loss in precision compared to other methods and is easier to compute than its jackknifed counterpart.

Next, we shall explore the role of MIs (and their possible associated errors) in mutagenicity assessment in different studies. Please note that we are not claiming that, neither experimental nor analytical, misconduct was indeed committed in any of the examples given below. We are merely stating and discussing the immediate issues that may arise whenever one chooses to make conclusions based solely on these simplistic measures (i.e., the MI and its modified versions).

Rather recently, Gupta et al. (2014) evaluated the sensitivity of three strains of S. typhimurium to hospital wastewater. They took a different approach to obtain mean MIs and standard deviations, by averaging the observed values among three independent experiments, each with three replicates. Such a procedure may have its own statistical issues, although our focus here is simply discussing the mutagenic effects they found for TA98 and TA100 strains. Despite the weakly positive mutagenicity, they recommend the continuous assessment of Jaipur and Delhi hospital wastewaters by the AT, since they potentially cause both frame-shift and base pair substitution mutations. In this case, only relatively minor, economic and human resources are of concern, since the most dangerous problem would be actually reporting a false negative (similar to a type-I error), which is less likely than a false positive, according to our results.

Table 1: Interval estimation of different mutagenicity indices, according to TA98 revertant count in each dose of quinoline (μg/plate), and their respective bootstrap and jackknife versions. Number in parenthesis $(\cdot,\cdot)$ denote the lower and upper bounds of each 95% confidence interval.

Dose	Index	Asy-N	Ext-B	Asy-B	Jac-N	Boo-N	Boo-P
10	MI	(0.62,1.07)	*	*	(0.25,1.38)	(0.53,1.13)	(0.61,1.22)
10	MI_m	(0.36,1.46)	*	*	(0.36,1.46)	(0.61,1.22)	(0.62,1.22)
10	$\tilde{MI}$	*	**	**	(0.32,1.44)	(0.37,1.28)	(0.55,1.40)
10	$\tilde{MI}_m$	*	(0.63,1.19)	(0.62,1.20)	(0.00,1.69)	(0.50,1.12)	(0.62,1.22)
33	MI	(0.38,1.93)	*	*	(0.34,1.88)	(0.60,1.64)	(0.73,1.80)
33	MI_m	(0.23,2.25)	*	*	(0.23,2.25)	(0.69,1.77)	(0.75,1.85)
33	$\tilde{MI}$	*	**	**	(0.47,2.07)	(0.41,2.05)	(0.55,2.20)
33	$\tilde{MI}_m$	*	(0.77,1.72)	(0.76,1.73)	(0.40,1.75)	(0.49,1.57)	(0.75,1.85)
100	MI	(0.47,3.47)	*	*	(0.58,3.22)	(0.99,2.89)	(1.20,3.16)
100	MI_m	(0.26,3.97)	*	*	(0.26,3.97)	(1.16,3.11)	(1.23,3.24)
100	$\tilde{MI}$	*	**	**	(0.73,3.27)	(0.36,3.37)	(0.93,4.00)
100	$\tilde{MI}_m$	*	(1.30,2.85)	(1.29,2.86)	(0.98,3.09)	(0.84,2.79)	(1.23,3.24)
333	MI	(1.36,2.09)	*	*	(0.50,2.81)	(1.10,2.28)	(1.29,2.49)
333	MI_m	(0.26,2.94)	*	*	(0.77,2.94)	(1.26,2.46)	(1.29,2.49)
333	$\tilde{MI}$	*	**	**	(0.68,3.03)	(0.90,2.69)	(1.14,2.73)
333	$\tilde{MI}_m$	*	(1.32,2.51)	(1.31,2.53)	(0.52,3.52)	(1.17,2.38)	(1.29,2.49)
1000	MI	(0.35,2.39)	*	*	(0.40,2.24)	(0.67,2.01)	(0.85,2.14)
1000	MI_m	(0.20,2.74)	*	*	(0.20,2.74)	(0.79,2.17)	(0.87,2.23)
1000	$\tilde{MI}$	*	**	**	(0.48,2.15)	(0.12,2.19)	(0.69,2.80)
1000	$\tilde{MI}_m$	*	(0.93,1.98)	(0.93,2.00)	(0.54,1.83)	(0.51,1.89)	(0.87,2.23)

Asy-N: asymptotic normal; Ext-B: exact binomial; Asy-B: asymptotic binomial; Jac-N: jackknifed normal; Boo-N: bootstrapped normal; Boo-P: percentile-based bootstrap interval. * method is not suited for the particular index; ** sample size is not enough to apply the method.

One should be aware that false positives are not always the lesser problem, even though most researchers are used to adopting the traditional 0.05 nominal level of significance (i.e., the probability of committing a type-I error and letting the probability of type-II error vary freely). Takemura et al. (2010) evaluated the pyrogallol-induced mutagenicity in *S. typhimurium* strains and they found lower MIs in the plates treated with antioxidants. Thus, they conclude that substances such as ascorbic acid (AA) present a protection effect against the hair dye ingredient (i.e., pyrogallol). Since the MIs they found in plates containing only pyrogallol range from 2.1 to 3.8, in average for TA100, and the reduction of revertants in the lowest concentration of AA is roughly 60%, then the protection of said antioxidant is ranging from 1.26 to 2.28 in terms of MIs. Depending on the fold rule adopted, these measures imply negative mutagenicity effect (which, in turn, corroborate their conclusion), even though such values can be expected to happen by chance alone in 95% of similar experiments, as our results show. In this case, a false positive may have been reported, analogous to a type-II error, further raising human health risk awareness.

Although varying a evaluating different sample sizes were not our objective, we must emphasize that protocol-recommended samples sizes are too small so that all the simplistic measures we took into consideration in this study can hardly be trustworthy from a statistical standpoint – which enlightens, though does not justify, the reasons behind the tedious, widespread argument that statistical analysis of AT data is either negligible or unnecessary. Indeed, as pointed out by several biostatisticians (Bernstein et al. 1982; Breslow 1984; Butturi-Gomes 2015; Hamada et al. 1994; Kim and Margolin 1999; Krewski and Franklin 1991; Krewski et al. 1993; Margolin et al. 1981; Myers et al. 1981; Stead et al. 1981), model-based measures and/or model fitting (e.g., regression analysis) are likely to be the better tools of choice whenever analyzing AT data.

In conclusion, the proposed indices had similar point and interval perfomances compared to the usual one, although the jackknifed MI_m and the bootstraped MI_m were slightly better overall. Furthermore, overdispersion showed to be an important factor whilst estimating mutagenic indices, diminishing their predictive capabilities.

Acknowledgments

The authors thank Dr. Thaís C. C. Fernandes and Dr. Maria Aparecida Marin-Morales for important biological insights and considerations, and CAPES for the financial support.

References

Akhtar MF et al. (2016) Toxicity appraisal of untreated dyeing industry wastewater based on chemical characterization and short term bioassays. Bulletin of Environmental Contamination and Toxicology 96:502-507 doi:10.1007/s00128-016-1759-x.

Ames BN, Durston WE, Yamasaki E, Lee FD (1973a) Carcinogens are mutagens: a simple test system combining liver homogenates for activation and bacteria for detection Proceedings of the National Academy of Sciences of the United States of America 70:2281-2285.

Ames BN, Lee FD, Durston WE (1973b) An improved bacterial test system for the detection and classification of mutagens and carcinogens. Proceedings of the National Academy of Sciences of the United States of America 70:782-786.

Barbee GC, Brown KW, Thomas JC, Donnelly KC, Murray HE (1996) Mutagenic activity (Ames test) of wood-preserving waste sludge applied to soil. Bulletin of Environmental Contamination and Toxicology 57:54-62 doi:10.1007/s001289900155.

Bernstein L, Kaldor J, McCann J, Pike MC (1982) An empirical approach to the statistical analysis of mutagenesis data from the *Salmonella* test. Mutation Research 97:267-281.

Breslow NE (1984) Extra-Poisson Variation in Log-Linear Models. Journal of the Royal Statistical Society: Series C 33:38-44.

Butturi-Gomes D (2015) Métodos estatísticos aplicados ao teste de *Salmonella*/mi crossoma: modelos, seleção e suas implicações. Escola Superior de Agricultura "Luiz de Queiroz", Universidade de São Paulo.

Cariello NF, Piegorsch WW (1996) The Ames test: the two-fold rule revisited. Mutation Research 369:23-31 doi:10.1016/S0165-1218(96)90044-0.

Edler L (1992) Statistical methods for short-term tests in genetic toxicology: the first fifteen years. Mutation Research 227:11-33 doi:10.1016/0165-1110(92)90024-4.

Efron B, Tibshirani RJ (1993) An introduction to the bootstrap. Chapman & Hall, Boca Raton.

Gupta P, Mathur P, Mathur N, Aarya B (2014) A comparative study of the sensitivities of *Salmonella* typhimurium strains TA98, TA100 and TA102 to hospital waste waters. Bulletin of Environmental Contamination and Toxicology 93:95-100 doi:10.1007/s00128-014-1274-x.

Hamada C, Wada T, Sakamoto Y (1994) Statistical Characterization of Negative Control Data in the Ames *Salmonella*/microsome test Environmental Health Perspectives 102:115-119

Hinde J, Demétrio CGB (1998) Overdispersion: models and estimation. Computational Statistics & Data Analysis 27:151-170.

Katz AJ (1978) Design and analysis of experiments on mutagenicity - I. Minimal sample sizes. Mutation Researh 50:301-307.

Katz AJ (1979) Design and analysis of experiments on mutagenicity - II. Analysis involving microorganisms. Mutation Researh 64:61-77.

Kim BS, Margolin BH (1999) Statistical methods of the Ames Salmonella assay: a review. Mutation Research 436:113-122.

Krewski D, Franklin C (eds) (1991) Statistics in toxicology. Gordon and Breach, New York.

Krewski D, Leroux BG, Bleur SR, Broekhaven LH (1993) Modelling the Ames *Salmonella*/microsome assay. Biometrics 49:499-510.

Lawless JF (1987) Negative binomial and mixed Poisson regression. The Canadian Journal of Statistics 15:209-225.

López LC, Bassi MD, Moretton J (1999) Influence of river water in the detection of Cr(VI) mutagenicity by the Ames test. Bulletin of Environmental Contamination and Toxicology 63:438-443 doi:10.1007/s001289900999.

Margolin BH, Kaplan N, Zeiger E (1981) Statistical analysis of the Ames *Salmonella*/microsome test. Proceedings of the National Academy of Sciences of the United States of America 78:3779-3783.

Molla DT, Muniswamy B (2012) Power of tests for overdispersion parameter in negative binomial regression model International Organization of Scientific Research Journal of Mathematics 1:29-36.

Mortelmans K, Zeiger E (2000) The Ames *Salmonella*/microsome mutagenicity assay. Mutation Research 455:29-60.

Myers LE, Saxton NH, Southerland LI, Wolff TJ (1981) Regression analysis of Ames test data. Environmental and Molecular Mutagenesis 3:575-586.

OECD (1997) Test No. 471: Bacterial Reverse Mutation Test. Washington, OECD Publishing. <http://www.oecd.org/chemicalsafety/risk-assessment/1948418.pdf>. Accessed 09-19-2016

R CORE TEAM (2015) R: A Language and Environment for Statistical Computing, 3.1 edn., Vienna: R Foundation for Statistical Computing.

Romero J, Ventura F, Caixach J, Rivera J, Guerrero R (1997) Identification and quantification of the mutagenic compound 3-chloro-4-(dichloromethyl)-5-hydroxy-2(5H)-furanone (MX) in chlorine-treated water. Bulletin of Environmental Contamination and Toxicology 59:715-722 doi:10.1007/s001289900539.

Stead AG, Hasselblad V, Creason JP, Claxton L (1981) Modelling the Ames test. Mutation Research 85:13-27

Takemura Y et al. (2010) Evaluation of pyrogallol-induced cytotoxicity in catalase-mutant *Escherichia coli* and mutagenicity in *Salmonella typhimurium*. Bulletin of Environmental Contamination and Toxicology 84:347-350 doi:10.1007/s00128-010-9936-9.

Umbuzeiro GA, Vargas VMF, Felzenszwalb I, Henriques JAP, Varanda E (2009) Teste de mutação reversa com *Salmonella typhimurium* (Teste de Ames, Ensaio *Salmonella*/microssoma). <http://mutagen-brasil.org.br/_img/_documentos/8abf3bb568d1c7ba57cd3fff7a22880c.pdf>. Accessed 09-22-2016.

Zar JH (2010) Biostatistical Analysis. 5th edn. Pearson Prentice Hall, Upper Saddle River.