



**Optimal inference via confidence distributions for two-by-two tables
modelled as Poisson pairs: fixed and random effects**

Céline Cunen*

University of Oslo, Oslo, Norway - cmlcunen@math.uio.no

Nils Lid Hjort

University of Oslo, Oslo, Norway - nils@math.uio.no

Abstract

This paper presents methods for meta-analysis of 2×2 tables, both with and without allowing heterogeneity in the treatment effects. Meta-analysis is common in medical research, but most existing methods are unsuited for 2×2 tables with rare events. Usually the tables are modelled as pairs of binomial variables, but we will model them as Poisson pairs. The methods presented here are based on confidence distributions, and offer optimal inference for the treatment effect parameter. We also propose an optimal method for inference on the ratio between treatment effects, and illustrate our methods on a real dataset.

Keywords: combining information; meta-analysis; rare events.

1. Introduction

Meta-analysis is a class of widely used methods for combination of information from independent studies. In medicine, meta-analyses of 2×2 tables are especially common. There are for example many studies investigating whether a certain drug is related to harmful side-effects. The usual set-up is to divide patients into a treatment and control group, and count the number of side-effect events in both groups. If the side-effect in question is very rare, individual studies can have low power, and meta-analysis is therefore especially important (Liu et al., 2014). The most common methods for meta-analysis of 2×2 tables rely on large sample approximations, which makes them unsuited for analysis of tables with rare events (Bradburn et al., 2007). This applies for example to the well-known Mantel-Haenszel method, the Peto method and the DerSimonian and Laird method (Bradburn et al. (2007) and Liu et al. (2014)). Recently Liu et al. (2014) introduced certain confidence distributions (CDs) based methods for this type of analysis. In this paper we present different CD-based methodology. Our methods utilise certain optimality theorems which in particular lead to CDs with a uniformly higher confidence power than for CDs constructed via the methods of Liu et al. (2014).

The number of events (cell-counts) in k 2×2 tables are often modelled as pairs of independent binomial random variables $(Y_{0,i}, Y_{1,i})$, with sample sizes $(m_{0,i}, m_{1,i})$ and event probabilities $(p_{0,i}, p_{1,i})$. Here $Y_{0,i}$ is the number of events among the control group in study i and $Y_{1,i}$ is the number of events in the treatment group in study i . Usually, the primary interest is not in the event probabilities themselves, but rather in the potential discrepancy between the event probabilities in the two groups. Whereas such a discrepancy can be defined in different fashions, the usual way in biostatistics is via the logistic transform:

$$p_{0,i} = \frac{e^{\theta_i}}{1 + e^{\theta_i}} \quad \text{and} \quad p_{1,i} = \frac{e^{\theta_i + \psi_i}}{1 + e^{\theta_i + \psi_i}} .$$

In particular, the odds ratio is

$$\text{OR}_i = \frac{p_{1,i}/(1 - p_{1,i})}{p_{0,i}/(1 - p_{0,i})} = \frac{e^{\theta_i + \psi_i}}{e^{\theta_i}} = e^{\psi_i} .$$

In the case of small event probabilities, the cell-counts can alternatively be modelled as pairs of independent Poisson variables,

$$Y_{0,i} \sim \text{Pois}(e_{0,i} \lambda_i) \quad \text{and} \quad Y_{1,i} \sim \text{Pois}(e_{1,i} \lambda_i \gamma_i) . \tag{1}$$

The argument for this is that binomial variables converge in distribution to Poisson when the sample size is large and the event probabilities are small. Here $e_{0,i}$ and $e_{1,i}$ are known exposure weights reflecting the sample sizes, i.e. we will expect more cases from a larger study, than a smaller one, independent of the event probabilities. We choose

$$e_{0,i} = m_{0,i}/100 \quad \text{and} \quad e_{1,i} = m_{1,i}/100.$$

Thus, λ_i can be interpreted as the risk of event per 100 patients in the control group, and γ_i as the multiplicative parameter associated with the potentially different risk (per 100 patients) for patients in the treatment group compared to the control group (henceforth called the treatment effect). Inference concerning odds ratio in the binomial model and γ_i in the Poisson model is similar when event probabilities are small and samples sizes are large. In fact,

$$\text{OR}_i = \frac{p_{1,i}/(1-p_{1,i})}{p_{0,i}/(1-p_{0,i})} \approx \frac{p_{1,i}}{p_{0,i}} \approx \frac{e_{1,i}\lambda_i\gamma_i/m_{1,i}}{e_{0,i}\lambda_i/m_{0,i}} = \frac{\lambda_i\gamma_i}{\lambda_i} = \gamma_i.$$

The λ_i parameters vary from table to table. In our paper we shall consider the fixed effect model, where the γ_i are equal to a common γ , as well as random effect models, where the γ_i are allowed to differ.

Our modus of inference will be that of CDs. A CD $C(\theta)$ for a parameter θ is a ‘distribution estimator’, i.e. a sample-dependent cumulative distribution function over the parameter space summarising all aspects of frequentist inference for θ (for a thorough introduction to inference with CDs see Schweder & Hjort (2015) and Xie & Singh (2013)). In addition to presenting the results as a cumulative distribution function, it is fruitful to represent the distribution of confidence as a confidence curve, defined as

$$\text{cc}(\theta) = |1 - 2C(\theta)|.$$

The confidence curve is a funnel plot pointing to the median confidence estimate $\hat{\theta} = C^{-1}(\frac{1}{2})$; also, confidence intervals for all levels α can be read off, see for example Figure 1.

For parameters from models of the exponential class, there exist *optimal* confidence distributions (Schweder & Hjort, 2015, Ch. 5). These are uniformly most powerful, in the sense of having lower expected confidence loss than all other CDs for a large class of penalty functions and all parameter values (see Schweder & Hjort (2002, 2015)). For both the binomial and Poisson model these optimal CDs are given in Schweder & Hjort (2015, Ch. 14). If the optimal (or other exact) CDs are hard or time-consuming to compute, approximate CDs can be constructed. There are several methods for this, and we will present some examples in section 3. In section 2 the optimal CD for the fixed effect Poisson model will be provided, while the Poisson model with heterogeneous treatment effects (random effects) is treated in section 4. Section 5 offers some concluding remarks.

2. Optimal confidence distribution for the fixed effect model

The Poisson fixed effect model has the following log-likelihood, writing $z_i = y_{0,i} + y_{1,i}$,

$$\ell(\gamma, \lambda_1, \dots, \lambda_k) = \sum_{i=1}^k \{y_{1,i} \log \gamma + z_i \log \lambda_i - (e_{0,i} + e_{1,i}\gamma)\lambda_i\}. \quad (2)$$

This is a log-likelihood of the exponential family class. By theorems in Schweder & Hjort (2015, Ch. 5) optimal inference for γ is based on its sufficient statistic ($\sum Y_{1,i}$) given the ancillary statistics (Z_1, \dots, Z_k). The conditional distribution of $Y_{1,i} | Z_i$ is seen to be binomial,

$$Y_{1,i} | (Z_i = z_i) \sim \text{Bin} \left(z_i, \frac{e_{1,i}\gamma}{e_{0,i} + e_{1,i}\gamma} \right). \quad (3)$$

From this distribution we can construct CDs for γ from each study (with half-correction for discreteness)

$$C_i(\gamma) = P_\gamma(Y_{1,i} > Y_{1,i,\text{obs}} | z_{i,\text{obs}}) + \frac{1}{2}P_\gamma(Y_{1,i} = Y_{1,i,\text{obs}} | z_{i,\text{obs}}). \quad (4)$$

Similarly, we can construct a combined CD for γ for all the studies, which is uniformly most powerful according to Schweder & Hjort (2015, Ch. 5),

$$C^*(\gamma) = P_\gamma(B > B_{\text{obs}} | z_{1,\text{obs}}, \dots, z_{k,\text{obs}}) + \frac{1}{2}P_\gamma(B = B_{\text{obs}} | z_{1,\text{obs}}, \dots, z_{k,\text{obs}}).$$

where $B = \sum_{i=1}^k Y_{1,i}$, and B_{obs} is the observed value of B . This is a sum of binomially distributed variables with different probability parameters ($e_{1,i}\gamma/(e_{0,i} + e_{1,i}\gamma)$), and its distribution might easily be simulated. As an illustration of this method, we have applied it to a medical datasets with six studies investigating death rates among heart attack patients (Normand, 1999). The treatment group received the drug Lidocaine and the control group did not. The sample sizes and number of deaths in the six studies are listed in Table 1.

Study	Sample size control	Events control	Sample size treatment	Events treatment
1	39	1	43	2
2	44	4	44	4
3	107	4	110	6
4	103	5	100	7
5	110	3	106	7
6	154	4	146	11

Table 1: Sample sizes and number of deaths for the six Lidocaine studies.

For the Lidocaine data, the combined CD for γ is shown in Figure 1, along with CDs from each study.

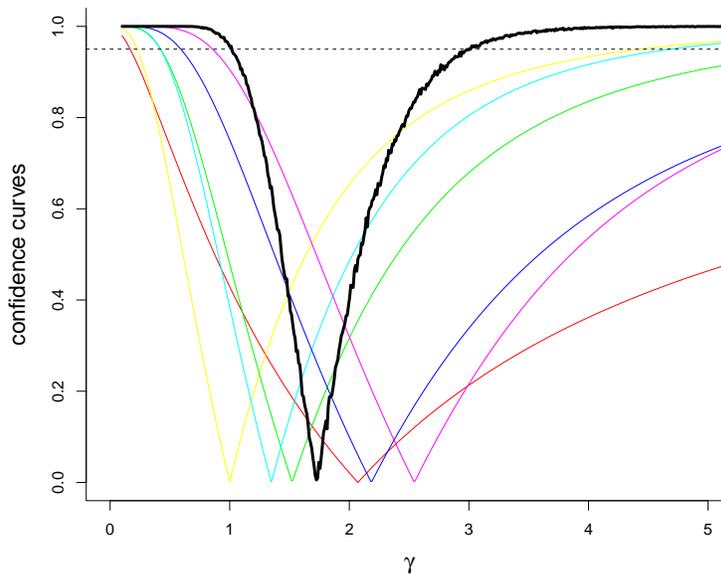


Figure 1: Confidence curves for γ for six Lidocaine studies (in colour) and combined confidence curve for all the studies (in black). A 95% confidence interval for γ can be read off (dashed line), [1.01, 3.01].

3. Approximate confidence distributions for the fixed effects model

The optimal CD can be slightly time-consuming to compute, but we can construct approximate CDs, offering a fast and often accurate alternative. Consider the log-likelihood function of the conditional model following from (3). We learn that the maximum conditional likelihood (MCL) estimator is a monotone function of $B = \sum Y_{1,i}$. Since CDs are invariant with respect to monotone transformations, the optimal CD in the previous section is actually the same as the CD based on the MCL estimator $\hat{\gamma}_{\text{mcl}}$. Ignoring half-corrections, we get

$$C^*(\gamma) = P_{\gamma}(B > B_{\text{obs}} \mid z_{1,\text{obs}}, \dots, z_{k,\text{obs}}) = P_{\gamma}(\hat{\gamma}_{\text{mcl}} > \hat{\gamma}_{\text{mcl,obs}}).$$

For the Poisson model the MCL estimator of γ is the solution of the following equation,

$$\sum_{i=1}^k \frac{y_{1,i}(e_{0,i} + e_{1,i}\gamma) - z_i e_{1,i}\gamma}{\gamma(e_{0,i} + e_{1,i}\gamma)} = 0.$$

Incidentally, the solution to the equation above is also the maximum likelihood (ML) estimator. Thus the unconditional and conditional likelihoods lead to the same estimators for this Poisson model. This pans out differently for other models, as for example the model of binomial pairs (Breslow, 1981).

Classical large samples methods for ML give us an approximate CD for γ , $C_{\text{approx}}(\gamma)$. The distribution of $\hat{\gamma}_{\text{mcl}}$ is approximately a normal (γ, \hat{J}^{-1}) , with \hat{J} the Hessian matrix associated with the log-likelihood function. This leads to the following approximate CD,

$$C_{\text{approx}}(\gamma) = \Phi((\gamma - \hat{\gamma}_{\text{mcl}})/\sqrt{\hat{J}^{-1}}).$$

Alternatively an approximate confidence distributions for γ can be constructed by χ_1^2 -transforming the deviance function of the profiled likelihood (see Schweder & Hjort (2015, Ch. 2-3)). The $k + 1$ -dimensional log-likelihood function in (2) is profiled by maximising over $\lambda_1, \dots, \lambda_k$, leading (in this case) to an explicit profile log-likelihood

$$\ell_{\text{prof}}(\gamma) = \sum_{i=1}^k \{y_{1,i} \log \gamma - z_i \log(e_{0,i} + e_{1,i}\gamma)\}.$$

The corresponding deviance function and the resulting confidence curve are

$$D(\gamma) = 2\{\ell_{\text{prof}}(\hat{\gamma}_{\text{ml}}) - \ell_{\text{prof}}(\gamma)\} \quad \text{and} \quad \text{cc}_a(\gamma) = \Gamma_1(D(\gamma)),$$

where $\Gamma_1()$ is the cumulative distribution function of the χ_1^2 . Both approximations work reasonably well, especially when there are many studies (large k), see Figure 2.

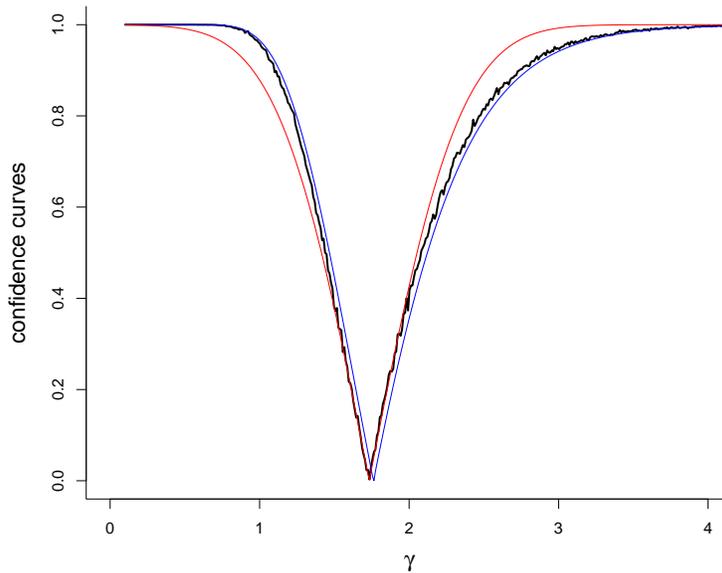


Figure 2: Optimal confidence curve for γ (in black) for the Lidocaine data, confidence curve based on the approximate distribution of the MCL (in red) and confidence curve based on profile deviance (in blue).

4. Optimal and approximate inference with heterogeneous treatment effects

In meta-analysis of 2×2 tables, the odds ratio from the binomial model is sometimes treated as a random effect (for example in the DerSimonian and Laird method). Similarly, cell-counts from 2×2 tables can be modelled with a Poisson model with a non-constant γ parameter. This allows the treatment effect γ_i to vary, as in (1). There are several possibilities regarding such models.

Our first suggestion is based on pairwise comparisons. Consider two studies at the time, for example study 1 and 2, with $Y_{0,1} \sim \text{Pois}(e_{0,1}\lambda_1)$, $Y_{1,1} \sim \text{Pois}(e_{1,1}\lambda_1\gamma_1)$ and with $Y_{0,2} \sim \text{Pois}(e_{0,2}\lambda_2)$, $Y_{1,2} \sim \text{Pois}(e_{1,2}\lambda_2\gamma_2)$.

We can construct a CD for the ratio between the two treatment effects, $\delta = \gamma_2/\gamma_1$. By arguments similar to those worked with in section 2, a power optimal CD for δ is found. It takes the form (with half-correction)

$$C_{1,2}(\delta) = P_\delta(Y_{1,2} > Y_{1,2,\text{obs}} | w_{\text{obs}}, z_{1,\text{obs}}, z_{2,\text{obs}}) + \frac{1}{2}P_\delta(Y_{1,2} = Y_{1,2,\text{obs}} | w_{\text{obs}}, z_{1,\text{obs}}, z_{2,\text{obs}}),$$

where $W = Y_{1,1} + Y_{1,2}$, $Z_1 = Y_{0,1} + Y_{1,1}$ and $Z_2 = Y_{0,2} + Y_{1,2}$. Thus we need the distribution of $Y_{1,2} | (W, Z_1, Z_2)$, and we find that it is an eccentric hypergeometric distribution, which may be expressed as

$$f(y_{1,2} | (w, z_1, z_2)) = \frac{\binom{z_1}{w-y_{1,2}} \binom{z_2}{y_{1,2}} \left(\frac{e_{0,1}e_{1,2}}{e_{1,1}e_{0,2}}\right)^{y_{1,2}} \delta^{y_{1,2}}}{\sum_{u=0}^w \binom{z_1}{w-u} \binom{z_2}{u} \left(\frac{e_{0,1}e_{1,2}}{e_{1,1}e_{0,2}}\right)^u \delta^u}, \text{ for } y_{1,2} \text{ from } 0 \text{ to } \min(w, z_2).$$

In this way we can construct CDs for $\delta_{i,j}$ for all pairs among the six Lidocaine studies. In Figure 3 the CD for $\delta_{2,6}$ is displayed. These are the two studies with the most different median confidence estimates for γ (see Figure 1). The CD for this δ indicates that 1 is a likely value (the 95% confidence interval is $[0.37, 17.81]$). Hence the γ s in the Lidocaine data do not seem sufficiently different to justify the use of a random effects model.

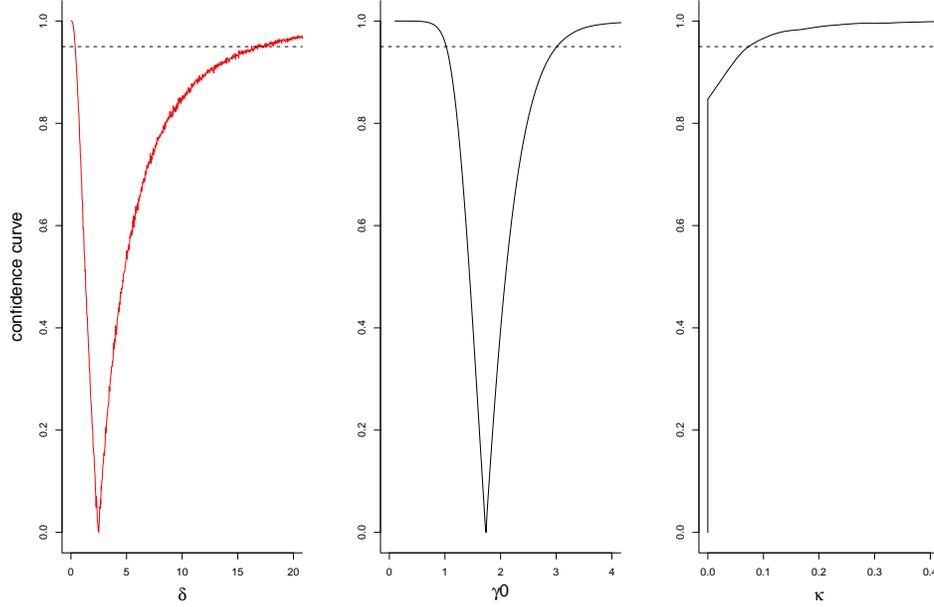


Figure 3: Three confidence curves are shown. In panel 1, for δ between the two most different Lidocaine studies (in red); in panel 2 for γ_0 ; in panel 3 for κ . The 95% confidence intervals can be read off, $[0.37, 17.81]$ for δ , $[1.02, 3.00]$ for γ_0 and $[0, 0.072]$ for κ .

The variation among the treatment effects from different studies can also be investigated in other ways. A natural suggestion is to model the variation of the γ_i with a beta distribution for the probabilities π_i from the conditional distribution of $Y_{1,i} | (Z_i = z_i)$,

$$Y_{1,i} | (Z_i = z_i, \gamma_i) \sim \text{Bin}(z_i, \pi_i) \quad \text{with} \quad \pi_i = \frac{e_{1,i}\gamma_i}{e_{0,i} + e_{1,i}\gamma_i},$$

where in addition $\pi_i \sim \text{Beta}\{\tau\pi_{0,i}, \tau(1 - \pi_{0,i})\}$ and $\pi_{0,i} = e_{1,i}\gamma_0 / (e_{0,i} + e_{1,i}\gamma_0)$. Letting $\kappa = (\tau + 1)^{-1}$, we have

$$E(\pi_i) = \pi_{0,i} \quad \text{and} \quad \text{Var}(\pi_i) = \kappa\pi_{0,i}(1 - \pi_{0,i}).$$

If $\tau = \infty$, then κ is equal to zero and we are back in the Poisson model with a common γ . We can construct approximate confidence curves for γ_0 by using profile deviance methods as described in section 3,

$$cc(\gamma_0) = \Gamma_1(D_1(\gamma_0)), \quad D_1(\gamma_0) = 2\{\ell_{1,\text{prof}}(\hat{\gamma}_0) - \ell_{1,\text{prof}}(\gamma_0)\}.$$

For κ , $cc(\kappa) = |1 - 2P_\kappa(Q_{\min} \geq Q_{\min, \text{obs}})|$ is a valid confidence curve, since it is (by construction) nearly uniformly distributed at the true parameter value. We use the natural statistic

$$Q_{\min} = \min_{\gamma_0} \sum_{i=1}^k \frac{\{y_{1,i} - z_i f_i(\gamma_0)\}^2}{z_i f_i(\gamma_0) \{1 - f_i(\gamma_0)\}}, \text{ where } f_i(\gamma_0) = \frac{e_{1,i} \gamma_0}{e_{0,i} + e_{1,i} \gamma_0},$$

and calculate it for the observed data (finding $Q_{\min, \text{obs}}$) and for a number of $y_{1,i}$ vectors, simulated under different κ values, with fixed z_i and $\hat{\gamma}_0$.

The CD for γ_0 allows us to make inference about the increase in risk related to the the treatment in question. While the CD for κ is used to assess the variation in the treatment effect. If the estimate for κ is small and the CD for κ has a significant point mass at zero, then the variation among the treatment effect is small and we can choose to use the fixed effect model from section 2. Otherwise, the random effect model may be preferable. For the Lidocaine example, the CD estimate for κ is zero and the point mass at zero is 0.85 (Figure 3). This supports our conclusion for the δ analysis above, that there is little variation among the γ s for the Lidocaine studies, and that the fixed effect model is the most appropriate.

5. Concluding remarks

There are many other aspects of meta-analysis of 2×2 tables (in the CD framework) worth considering. First, we have applied our CD-methods to several other datasets, including a small dataset from Narum et al. (2014), with five studies investigating whether drugs of a certain type (corticosteroids) are related to gastrointestinal bleeding for patients in ambulatory care. These data poses some challenges for many traditional meta-analysis methods because most of the studies have zero event entries, either in the treatment group or the control group. Many existing methods demand 0.5 corrections to zero events in such situations, but these kinds of corrections can give large biases and are not recommended (Bradburn et al., 2007). One of the strengths of our method is that it can produce a combined confidence distribution in this case, without needing to add anything to the data. Secondly, we have implemented the alternative CD based combination method from Liu et al. (2014), and compared it to our methods. The two classes of methods can yield similar results, but ours yield stronger confidence. Lastly we have considered extensions to 2×3 tables (which the Poisson model can handle without problems) and alternative ways to model the heterogeneous treatment effects (for example with a Poisson-gamma model).

References

- Bradburn, M. J., Deeks, J. J., Berlin, J. A., & Russell Localio, A. (2007). Much ado about nothing: a comparison of the performance of meta-analytical methods with rare events. *Statistics in Medicine*, 26, 53–77.
- Breslow, N. (1981). Odds ratio estimators when the data are sparse. *Biometrika*, 68, 73–84.
- Liu, D., Liu, R. Y., & Xie, M. (2014). Exact meta-analysis approach for discrete data and its application to 2×2 tables with rare events. *Journal of the American Statistical Association*, 109, 1450–1465.
- Narum, S., Westergren, T., & Klemp, M. (2014). Corticosteroids and risk of gastrointestinal bleeding: a systematic review and meta-analysis. *BMJ open*, 4.
- Normand, S.-L. T. (1999). Tutorial in biostatistics meta-analysis: formulating, evaluating, combining, and reporting. *Statistics in Medicine*, 18, 321–359.
- Schweder, T. & Hjort, N. L. (2002). Confidence and likelihood. *Scandinavian Journal of Statistics*, 29, 309–332.
- Schweder, T. & Hjort, N. L. (2015). *Confidence, Likelihood, Probability: Inference With Confidence Distributions*. Cambridge University Press.
- Xie, M. & Singh, K. (2013). Confidence distribution, the frequentist distribution estimator of a parameter: a review [with discussion and a rejoinder]. *International Statistical Review*, 81, 3–39.