Supplement to the *Discussion of "CARS: Covariate assisted ranking and screening for large-scale two-sample inference" by Cai et al.*

Guo Yu[*] Jacob Bien[†] Daniela Witten[‡]

December 18, 2018

We adapt the notation introduced in the discussion. Furthermore, suppose the random errors

$$\begin{pmatrix} \varepsilon_{ij}(\ell) \\ \tilde{\varepsilon}_{i\sigma(j)}(\ell) \end{pmatrix} \sim N\left(\mathbf{0}, \begin{pmatrix} \nu_j^2(\ell) & \tau_{j\sigma(j)}(\ell) \\ \tau_{j\sigma(j)}(\ell) & \varsigma_{\sigma(j)}^2(\ell) \end{pmatrix}\right) \qquad \text{for} \quad j \in \{1, ..., m\},$$

and these are independent across j, ℓ and i. Recall that we assume the random errors are independent of the mean vectors $\boldsymbol{\mu}(\ell)$ and $\tilde{\boldsymbol{\mu}}(\ell)$.

1 Properties of primary and auxiliary statistics

Recall that the "primary statistic" for testing $\mu_j(1) = \mu_j(2)$ is

$$T_j = C_j \left(\bar{X}_j(1) - \bar{X}_j(2)\right)$$

for some standardizing constant C_j. We consider a pair of "auxiliary statistics,"

$$R_j = D_j \left(\bar{X}_j(1) + \kappa_j \bar{X}_j(2)\right), \qquad S_j = E_j \left(\bar{Z}_{\sigma(j)}(1) + \xi_j \bar{Z}_{\sigma(j)}(2)\right),$$

where D_j and E_j are some standardizing constants, and

$$\kappa_j = \frac{n_2 \nu_j^2(1)}{n_1 \nu_j^2(2)}, \qquad \xi_j = \frac{n_2 \tau_{j\sigma(j)}(1)}{n_1 \tau_{j\sigma(j)}(2)}. \tag{1}$$

The following result gives the key conditional independence property:

Proposition 1. *The primary statistic and the pair of auxiliary statistics satisfy*

$$\begin{aligned} &f(t_j, r_j, s_j | \mu_j(1), \mu_j(2), \tilde{\mu}_{\sigma(j)}(1), \tilde{\mu}_{\sigma(j)}(2)) \\ =&f(t_j | \mu_j(1), \mu_j(2), \tilde{\mu}_{\sigma(j)}(1), \tilde{\mu}_{\sigma(j)}(2)) f(r_j, s_j | \mu_j(1), \mu_j(2), \tilde{\mu}_{\sigma(j)}(1), \tilde{\mu}_{\sigma(j)}(2)). \end{aligned} \tag{2}$$

[*]Department of Statistics, University of Washington, Seattle, Washington, 98195, gy63@uw.edu

[†]Data Sciences and Operations, Marshall School of Business, University of Southern California, Los Angeles, CA 90089, jbien@usc.edu

[‡]Department of Statistics and Biostatistics, University of Washington, Seattle, Washington, 98195, dwitten@uw.edu

Proof. We first show that conditioned on $\mu_j(1), \mu_j(2), \tilde{\mu}_{\sigma(j)}(1), \tilde{\mu}_{\sigma(j)}(2)$, the statistic (T_j, R_j, S_j) is jointly normal. Denote $\boldsymbol{\varepsilon}(1) \in \mathbb{R}^{2n_1}$ as the vector of random errors by stacking $(\varepsilon_{ij}(1), \tilde{\varepsilon}_{i\sigma(j)}(1))^T$ for $i = 1, ..., n_1$, and similarly $\boldsymbol{\varepsilon}(2) \in \mathbb{R}^{2n_2}$ as the vector of random errors by stacking $(\varepsilon_{ij}(2), \tilde{\varepsilon}_{i\sigma(j)}(2))^T$ for $i = 1, ..., n_2$. Then $\boldsymbol{\varepsilon}(1)$ and $\boldsymbol{\varepsilon}(2)$ follow multivariate normal distributions. By the independence between $\boldsymbol{\varepsilon}(1)$ and $\boldsymbol{\varepsilon}(2)$, we have that $\boldsymbol{\varepsilon} = (\boldsymbol{\varepsilon}(1)^T, \boldsymbol{\varepsilon}(2)^T)^T \in \mathbb{R}^{2n_1+2n_2}$ follows a multivariate normal distribution. Now conditioned on $\mu_j(1), \mu_j(2), \tilde{\mu}_{\sigma(j)}(1), \tilde{\mu}_{\sigma(j)}(2)$, the vector (T_j, R_j, S_j) is a linear transformation of $\boldsymbol{\varepsilon}$. Thus by definition, (T_j, R_j, S_j) is jointly normal.

By (1), we have

$$
\begin{aligned}
&\operatorname{Cov}\left[(T_j, R_j, S_j)|\mu_j(1), \mu_j(2), \tilde{\mu}_{\sigma(j)}(1), \tilde{\mu}_{\sigma(j)}(2)\right] \\
=&\begin{pmatrix}
C_j^2\left(\frac{\nu_j^2(1)}{n_1} + \frac{\nu_j^2(2)}{n_2}\right) & 0 & 0 \\
0 & D_j^2\left(\frac{\nu_j^2(1)}{n_1} + \frac{n_2}{n_1^2}\frac{\nu_j^4(1)}{\nu_j^2(2)}\right) & D_jE_j\left(\frac{\tau_{j\sigma(j)}(1)}{n_1} + \frac{\nu_j^2(1)}{\nu_j^2(2)}\frac{n_2\tau_{j\sigma(j)}(1)}{n_1^2}\right) \\
0 & D_jE_j\left(\frac{\tau_{j\sigma(j)}(1)}{n_1} + \frac{\nu_j^2(1)}{\nu_j^2(2)}\frac{n_2\tau_{j\sigma(j)}(1)}{n_1^2}\right) & E_j^2\left(\frac{\varsigma_j^2(1)}{n_1} + \frac{\tau_{j\sigma(j)}^2(1)}{\tau_{j\sigma(j)}^2(2)}\frac{n_2\varsigma_j^2(2)}{n_1^2}\right)
\end{pmatrix}.
\end{aligned}
$$

Therefore, T_j is independent of (R_j, S_j) conditional on $\mu_j(1), \mu_j(2), \tilde{\mu}_{\sigma(j)}(1), \tilde{\mu}_{\sigma(j)}(2)$. □

Now we derive the key property that is needed to construct the CARS procedure with FDR control. Let $\theta_{1j} = \mathbb{1}\{\mu_j(1) \neq \mu_j(2)\}$, $\theta_{2j} = \mathbb{1}\{\mu_j(1) \neq 0 \text{ or } \mu_j(2) \neq 0\}$, and $\theta_{3\sigma(j)} = \mathbb{1}\{\tilde{\mu}_{\sigma(j)}(1) \neq 0 \text{ or } \tilde{\mu}_{\sigma(j)}(2) \neq 0\}$. By the hierarchy,

$$
\theta_{3\sigma(j)} = 0 \implies \theta_{2j} = 0 \implies \theta_{1j} = 0. \tag{3}
$$

And we have the following results.

Proposition 2. *The conditional independence property in* (2) *implies that*

$$
\begin{aligned}
f\left(t_j, r_j, s_j|\theta_{1j} = 0, \theta_{2j} = 0, \theta_{3\sigma(j)} = 0\right) &= f\left(t_j|\theta_{1j} = 0\right) f\left(r_j, s_j|\theta_{1j} = 0, \theta_{2j} = 0, \theta_{3\sigma(j)} = 0\right); \\
f\left(t_j, r_j, s_j|\theta_{1j} = 0, \theta_{2j} = 0, \theta_{3\sigma(j)} = 1\right) &= f\left(t_j|\theta_{1j} = 0\right) f\left(r_j, s_j|\theta_{1j} = 0, \theta_{2j} = 0, \theta_{3\sigma(j)} = 1\right); \\
f\left(t_j, r_j, s_j|\theta_{1j} = 0, \theta_{2j} = 1, \theta_{3\sigma(j)} = 1\right) &= f\left(t_j|\theta_{1j} = 0\right) f\left(r_j, s_j|\theta_{1j} = 0, \theta_{2j} = 1, \theta_{3\sigma(j)} = 1\right).
\end{aligned}
$$

Moreover, the primary and the auxiliary statistics are conditionally independent under the null, i.e.,

$$
f\left(t_j, r_j, s_j|\theta_{1j} = 0\right) = f\left(t_j|\theta_{1j} = 0\right) f\left(r_j, s_j|\theta_{1j} = 0\right). \tag{4}
$$

Proof. We adapt similar arguments in A.1 of Cai et al. (2019). First note that

$$
\begin{aligned}
&f\left(t_j, r_j, s_j|\theta_{1j} = 0, \theta_{2j} = 0, \theta_{3\sigma(j)} = 0\right) \\
=&f\left(t_j, r_j, s_j|\mu_j(1) = \mu_j(2) = \tilde{\mu}_{\sigma(j)}(1) = \tilde{\mu}_{\sigma(j)}(2) = 0\right) \\
=&f\left(t_j|\mu_j(1) = \mu_j(2) = \tilde{\mu}_{\sigma(j)}(1) = \tilde{\mu}_{\sigma(j)}(2) = 0\right) f\left(r_j, s_j|\mu_j(1) = \mu_j(2) = \tilde{\mu}_{\sigma(j)}(1) = \tilde{\mu}_{\sigma(j)}(2) = 0\right) \\
=&f\left(t_j|\theta_{1j} = 0, \mu_j(1) = \mu_j(2) = \tilde{\mu}_{\sigma(j)}(1) = \tilde{\mu}_{\sigma(j)}(2) = 0\right) f\left(r_j, s_j|\theta_{1j} = 0, \theta_{2j} = 0, \theta_{3\sigma(j)} = 0\right) \\
=&f\left(t_j|\theta_{1j} = 0\right) f\left(r_j, s_j|\theta_{1j} = 0, \theta_{2j} = 0, \theta_{3\sigma(j)} = 0\right).
\end{aligned}
$$

We apply (3) to get the first equality above, and (2) implies the second equality. The last equality holds because given $\theta_{1j} = 0$, t_j is a linear combination of $\varepsilon_{ij}(1)$ for $i = 1, ..., n_1$ and $\varepsilon_{ij}(2)$ for $i = 1, ..., n_2$, which is independent of $\mu_j(1)$, $\mu_j(2)$, $\tilde{\mu}_{\sigma(j)}(1)$, and $\tilde{\mu}_{\sigma(j)}(2)$. Similarly, given $\theta_{3\sigma(j)} = 0$, r_j and s_j are linear combinations of $\varepsilon_{ij}(1)$, $\tilde{\varepsilon}_{i\sigma(j)}(1)$ for $i = 1, ..., n_1$, and $\varepsilon_{ij}(2)$, $\tilde{\varepsilon}_{i\sigma(j)}(2)$ for $i = 1, ..., n_2$, which are independent of $\mu_j(1)$, $\mu_j(2)$, $\tilde{\mu}_{\sigma(j)}(1)$, and $\tilde{\mu}_{\sigma(j)}(2)$.

Denote $\mathcal{A} = \{(0, 0, y, z) : |y| + |z| \neq 0\}$, and F as the distribution function of $\boldsymbol{\gamma}_j = (\mu_j(1), \mu_j(2), \tilde{\mu}_{\sigma(j)}(1), \tilde{\mu}_{\sigma(j)}(2))$. Let $Q_{\mathcal{A}} = \Pr(\boldsymbol{\gamma}_j \in \mathcal{A})$. Then

$$
\begin{aligned}
&f(t_j, r_j, s_j|\theta_{1j} = 0, \theta_{2j} = 0, \theta_{3\sigma(j)} = 1) = f(t_j, r_j, s_j|\boldsymbol{\gamma} \in \mathcal{A})\\
=&Q_{\mathcal{A}}^{-1} \int_{\mathcal{A}} f(t_j, r_j, s_j|\boldsymbol{\gamma}_j = \boldsymbol{a})dF(\boldsymbol{a})\\
=&Q_{\mathcal{A}}^{-1} \int_{\mathcal{A}} f(t_j|\boldsymbol{\gamma}_j = \boldsymbol{a})f(r_j, s_j|\boldsymbol{\gamma}_j = \boldsymbol{a})dF(\boldsymbol{a})\\
=&Q_{\mathcal{A}}^{-1} \int_{\mathcal{A}} f(t_j|\boldsymbol{\gamma}_j = \boldsymbol{a}, \theta_{1j} = 0)f(r_j, s_j|\boldsymbol{\gamma}_j = \boldsymbol{a})dF(\boldsymbol{a})\\
=&f(t_j|\theta_{1j} = 0)Q_{\mathcal{A}}^{-1} \int_{\mathcal{A}} f(r_j, s_j|\boldsymbol{\gamma}_j = \boldsymbol{a})dF(\boldsymbol{a})\\
=&f(t_j|\theta_{1j} = 0)f(r_j, s_j|\boldsymbol{\gamma}_j \in \mathcal{A})\\
=&f(t_j|\theta_{1j} = 0)f(r_j, s_j|\theta_{1j} = 0, \theta_{2j} = 0, \theta_{3\sigma(j)} = 1).
\end{aligned}
$$

Denote $\mathcal{B} = \{(x, x, y, z) : x \neq 0, |y| + |z| \neq 0\}$, and $Q_{\mathcal{B}} = \Pr(\boldsymbol{\gamma}_j \in \mathcal{B})$. Then

$$
\begin{aligned}
&f(t_j, r_j, s_j|\theta_{1j} = 0, \theta_{2j} = 1, \theta_{3\sigma(j)} = 1) = f(t_j, r_j, s_j|\boldsymbol{\gamma}_j \in \mathcal{B})\\
=&Q_{\mathcal{B}}^{-1} \int_{\mathcal{B}} f(t_j, r_j, s_j|\boldsymbol{\gamma}_j = \boldsymbol{b})dF(\boldsymbol{b})\\
=&Q_{\mathcal{B}}^{-1} \int_{\mathcal{B}} f(t_j|\boldsymbol{\gamma}_j = \boldsymbol{b})f(r_j, s_j|\boldsymbol{\gamma}_j = \boldsymbol{b})dF(\boldsymbol{b})\\
=&Q_{\mathcal{B}}^{-1} \int_{\mathcal{B}} f(t_j|\boldsymbol{\gamma}_j = \boldsymbol{b}, \theta_{1j} = 0)f(r_j, s_j|\boldsymbol{\gamma}_j = \boldsymbol{b})dF(\boldsymbol{b})\\
=&f(t_j|\theta_{1j} = 0)Q_{\mathcal{B}}^{-1} \int_{\mathcal{B}} f(r_j, s_j|\boldsymbol{\gamma}_j = \boldsymbol{b})dF(\boldsymbol{b})\\
=&f(t_j|\theta_{1j} = 0)f(r_j, s_j|\boldsymbol{\gamma}_j \in \mathcal{B})\\
=&f(t_j|\theta_{1j} = 0)f(r_j, s_j|\theta_{1j} = 0, \theta_{2j} = 1, \theta_{3\sigma(j)} = 1).
\end{aligned}
$$

By the hierarchy (3), we have that $\Pr(\theta_{2j} = 1, \theta_{3\sigma(j)} = 0|\theta_{1j} = 0) = 0$. Combining these three results, we have

$$
\begin{aligned}
&f(t_j, r_j, s_j|\theta_{1j} = 0) = \sum_{(a,b)\in\{0,1\}^2} f(t_j, r_j, s_j|\theta_{1j} = 0, \theta_{2j} = a, \theta_{3\sigma(j)} = b)\Pr(\theta_{2j} = a, \theta_{3\sigma(j)} = b|\theta_{1j} = 0)\\
=&f\left(t_j|\theta_{1j} = 0\right)\left[f\left(r_j, s_j|\theta_{1j} = 0, \theta_{2j} = 0, \theta_{3\sigma(j)} = 0\right)\Pr(\theta_{2j} = 0, \theta_{3\sigma(j)} = 0|\theta_{1j} = 0)\right.\\
&\qquad + f(r_j, s_j|\theta_{1j} = 0, \theta_{2j} = 0, \theta_{3\sigma(j)} = 1)\Pr(\theta_{2j} = 0, \theta_{3\sigma(j)} = 1|\theta_{1j} = 0)\\
&\qquad \left. + f(r_j, s_j|\theta_{1j} = 0, \theta_{2j} = 1, \theta_{3\sigma(j)} = 1)\Pr(\theta_{2j} = 1, \theta_{3\sigma(j)} = 1|\theta_{1j} = 0)\right]\\
=&f\left(t_j|\theta_{1j} = 0\right)f(r_j, s_j|\theta_{1j} = 0).
\end{aligned}
$$

□

So we could simplify the joint distribution of (t_j, r_j, s_j) as

$$
\begin{aligned}
f(t_j, r_j, s_j) = & \sum_{(a,b,c)\in\{0,1\}^3} \pi_{abc} f(t_j, r_j, s_j|\theta_{1j} = a, \theta_{2j} = b, \theta_{3\sigma(j)} = c) \\
= & \pi_{000} f(t_j, r_j, s_j|\theta_{1j} = 0, \theta_{2j} = 0, \theta_{3\sigma(j)} = 0) \\
& + \pi_{001} f(t_j, r_j, s_j|\theta_{1j} = 0, \theta_{2j} = 0, \theta_{3\sigma(j)} = 1) \\
& + \pi_{011} f(t_j, r_j, s_j|\theta_{1j} = 0, \theta_{2j} = 1, \theta_{3\sigma(j)} = 1) \\
& + \pi_{111} f(t_j, r_j, s_j|\theta_{1j} = 1, \theta_{2j} = 1, \theta_{3\sigma(j)} = 1) \\
= & \pi_{000} f(t_j|\theta_{1j} = 0) f(r_j, s_j|\theta_{2j} = 0, \theta_{3\sigma(j)} = 0) \\
& + \pi_{001} f(t_j|\theta_{1j} = 0) f(r_j, s_j|\theta_{2j} = 0, \theta_{3\sigma(j)} = 1) \\
& + \pi_{011} f(t_j|\theta_{1j} = 0) f(r_j, s_j|\theta_{2j} = 1, \theta_{3\sigma(j)} = 1) \\
& + \pi_{111} f(t_j, r_j, s_j|\theta_{1j} = 1, \theta_{2j} = 1, \theta_{3\sigma(j)} = 1).
\end{aligned} \tag{5}
$$

2 Oracle procedure

Following Section 3 of Cai et al. (2019), we define the oracle statistic

$$
\begin{aligned}
T_{OR}^{(j)}(t_j, r_j, s_j) &= \Pr(\theta_{1j} = 0|T_j = t_j, R_j = r_j, S_j = s_j) \\
&= \frac{\Pr(\theta_{1j} = 0, T_j = t_j, R_j = r_j, S_j = s_j)}{f(t_j, r_j, s_j)} \\
&= \frac{f(t_j, r_j, s_j|\theta_{1j} = 0)\Pr(\theta_{1j} = 0)}{f(t_j, r_j, s_j)} \\
&= \frac{f(t_j|\theta_{1j} = 0) f(r_j, s_j|\theta_{1j} = 0)\Pr(\theta_{1j} = 0)}{f(t_j, r_j, s_j)}.
\end{aligned} \tag{6}
$$

The term $f(t_j|\theta_{1j} = 0)$ is known by the model specification, and the joint distribution $f(t_j, r_j, s_j)$ can be written as in (5). Approximating and estimating the term $f(r_j, s_j|\theta_{1j} = 0)\Pr(\theta_{1j} = 0) =: q^*(r_j, s_j)$ can be done in a similar manner as discussed in Sections 3.2 and 3.3 of Cai et al. (2019).

A careful examination of the proof of Theorem 1 in Cai et al. (2019) suggests that Theorem 1 in Cai et al. (2019) also holds here, i.e., the testing rule $\{\mathbb{1}(T_{OR}^{(j)} < \lambda) : 1 \le j \le m\}$ has mFDR bounded by λ, and there is no efficiency loss by utilizing the auxiliary statistics.

Theorem 3 (Theorem 1 in Cai et al. (2019))**.** *Consider the oracle statistic* (6)*. Then*

(a) For $0 < \lambda \le 1$, let $Q_{OR}(\lambda)$ be the mFDR level of testing rule $\left[\mathbb{1}\left\{T_{OR}^{(j)}(t_j, r_j, s_j) < \lambda\right\} : 1 \le j \le m\right]$. Then $Q_{OR}(\lambda) < \lambda$ and $Q_{OR}(\lambda)$ is non-decreasing in λ.

(b) Suppose we choose $\alpha < Q_{OR}(1)$. Then the oracle threshold $\lambda_{OR} = \sup\{\lambda : Q_{OR}(\lambda) \le \alpha\}$ exists uniquely and $Q_{OR}(\lambda_{OR}) = \alpha$. Furthermore, define the oracle rule

$$\boldsymbol{\delta}_{OR} = \left\{\mathbb{1}(T_{OR}^{(j)}(t_j, r_j, s_j) < \lambda_{OR}) : 1 \le j \le m\right\}.$$

Then $\boldsymbol{\delta}_{OR}$ is optimal in the sense that $ETP(\boldsymbol{\delta}) \le ETP(\boldsymbol{\delta}_{OR})$ for any testing rule $\boldsymbol{\delta}$ based on $\{T_j, R_j, S_j\}$ such that $mFDR(\boldsymbol{\delta}) \le \alpha$.

Proof. The proof of Theorem 1 in Cai et al. (2019) can be directly used here to prove Theorem 3. Here we show the proof of equality (7.1) in Cai et al. (2019), which is not trivial but was not given in the original paper.

First denote $Q_{OR}(t) = \alpha_t$. By the definition of mFDR, we have

$$\begin{aligned} Q_{OR}(t) &= \frac{\mathrm{E}\left[\sum_j (1-\theta_{1j})\mathbb{1}\{T_{OR}^{(j)} < t\}\right]}{\mathrm{E}\left[\sum_j \mathbb{1}\{T_{OR}^{(j)} < t\}\right]} \\ &= \frac{\mathrm{E}_{\boldsymbol{T},\boldsymbol{R},\boldsymbol{S}}\left\{\mathrm{E}\left[\sum_j (1-\theta_{1j})\mathbb{1}\{T_{OR}^{(j)} < t\}|\boldsymbol{T},\boldsymbol{R},\boldsymbol{S}\right]\right\}}{\mathrm{E}_{\boldsymbol{T},\boldsymbol{R},\boldsymbol{S}}\left\{\mathrm{E}\left[\sum_j \mathbb{1}\{T_{OR}^{(j)} < t\}|\boldsymbol{T},\boldsymbol{R},\boldsymbol{S}\right]\right\}} \\ &= \frac{\mathrm{E}_{\boldsymbol{T},\boldsymbol{R},\boldsymbol{S}}\left\{\sum_j \mathbb{1}\{T_{OR}^{(j)} < t\}\mathrm{E}\left[(1-\theta_{1j})|\boldsymbol{T},\boldsymbol{R},\boldsymbol{S}\right]\right\}}{\mathrm{E}_{\boldsymbol{T},\boldsymbol{R},\boldsymbol{S}}\left\{\sum_j \mathbb{1}\{T_{OR}^{(j)} < t\}\mathrm{E}\left[1|\boldsymbol{T},\boldsymbol{R},\boldsymbol{S}\right]\right\}} \\ &= \frac{\mathrm{E}_{\boldsymbol{T},\boldsymbol{R},\boldsymbol{S}}\left\{\sum_j \mathbb{1}\{T_{OR}^{(j)} < t\}T_{OR}^{(j)}\right\}}{\mathrm{E}_{\boldsymbol{T},\boldsymbol{R},\boldsymbol{S}}\left\{\sum_j \mathbb{1}\{T_{OR}^{(j)} < t\}\right\}}, \end{aligned}$$

where the last equality holds by the definition of $T_{OR}^{(j)}$,

$$T_{OR}^{(j)} = \Pr(\theta_{1j} = 0|T_j, R_j, S_j) = \mathrm{E}\left[(1-\theta_{1j})|T_j, R_j, S_j\right],$$

and the subscript of expectation $(\boldsymbol{T},\boldsymbol{R},\boldsymbol{S})$ indicates that the expectation is taken over the joint distribution of $(\boldsymbol{T},\boldsymbol{R},\boldsymbol{S})$. Therefore, $Q_{OR}(t) = \alpha_t$ implies

$$\mathrm{E}_{\boldsymbol{T},\boldsymbol{R},\boldsymbol{S}}\left\{\sum_j \mathbb{1}\{T_{OR}^{(j)} < t\}(T_{OR}^{(j)} - \alpha_t)\right\} = 0,$$

which is (7.1) in Cai et al. (2019). And the result of the theorem follows. □

References

Cai, T. T., Sun, W. & Wang, W. (2019), 'CARS: Covariate assisted ranking and screening for large-scale two-sample inference'.