

Some Applications of Nonparanormal Adjusted Marginal Inference

Susanne Dandl
Universität Zürich

Torsten Hothorn
Universität Zürich

Abstract

Covariate adjustment is recommended to improve precision of estimated treatment effects, even in randomised trials. For non-collapsible effect measures (such as Cohen's d , odds and hazard ratios), conditioning on covariates can alter the effect interpretation such that effects are not comparable given different adjustment sets. A novel nonparanormal adjusted marginal inference approach (NAMI) can be formulated for randomised controlled trials to derive marginal effects from the estimated joint distribution of covariates and outcomes. NAMI preserves the interpretation of treatment effects regardless of covariates used, while enhancing precision. We present applications of NAMI with continuous, binary and right-censored survival outcomes. Data is analysed by a proof-of-concept implementation of NAMI available in the **tram** add-on package to the R system for statistical computing.

Keywords: marginal effect, noncollapsibility, covariate adjustment, randomised trial, transformation model.

1. Introduction

We present three applications from different domains illustrating the potential to estimate treatment effects more precisely with nonparanormal adjusted marginal inference (NAMI, [Dandl and Hothorn 2025](#)): A toxicity study demonstrates equivalence of a continuous outcome by Cohen's d , an efficiency trial compares a binary outcome between two arms by means of an odds ratio, and a survival study measures differences in terms of a hazard ratio.

We apply the proof-of-concept implementation of nonparanormal adjusted marginal inference available in package **tram**. All results of this document can be reproduced from the NAMI demo available in the package:

```
R> install.packages("tram")  
R> demo("NAMI", package = "tram")
```

Please cite this document as: Susanne Dandl and Torsten Hothorn (2025) Some Applications of Nonparanormal Adjusted Marginal Inference. R package vignette version 1.3-0, URL <https://doi.org/10.32614/CRAN.package.tram>.

2. Continuous outcome: Immunotoxicity study on Chloramine

To assess the toxicity of Chloramine, a study on mice was conducted as part of the [National Toxicology Program \(2000\)](#). Female mice were randomly assigned to two groups of which one group received Chloramine-dosed water and the other not. Repeated measurements of weights were conducted on days 1, 8, 15, 22, and 29 in five dose groups (0, 2, 10, 20, 100 mg/kg).

The following data set is a subset of the data obtained from the [National Toxicology Program \(2000\)](#). It focuses on the comparison of the highest dose group ($W = 1, N_1 = 40$) with the control group ($W = 0, N_0 = 40$) with respect to the weight at day 29 (denoted as outcome Y). Information on the weight on day 1 is also available and is used for adjustment, in the following.

The research question of “no effect of Chloramine on weight” is formulated in terms of the hypotheses $H_0 : |\tau| \geq \delta$ and $H_1 : |\tau| < \delta$ where Cohen’s d is the marginal treatment effect τ . H_0 is rejected at 5% when the 95% confidence interval for τ is completely contained in the equivalence interval $(-\delta, \delta)$. We set $\delta = 0.36$ as recommended by [Wellek \(2010\)](#) (Table 1.1.), acknowledging this as a potential oversimplification.

2.1. Unadjusted marginal inference

The following fits an unadjusted marginal outcome model featuring Cohen’s d :

```
R> m0 <- Lm(y ~ w, data = immun)

R> coef(m0)                                     ### marginal Cohen's d

      w100
0.04817

R> sqrt(vcov(m0))                             ### observed

      w100
w100 0.2236

R> sqrt(2/nrow(immun) * (coef(m0)^2/4 + 2)) ### expected

      w100
0.2236

R> confint(m0)                                 ### Wald

      2.5 % 97.5 %
w100 -0.3902 0.4865
```

The unadjusted estimate of Cohen’s d is $\hat{\tau} = 0.048$. The standard error based on the observed Fisher information is $SE(\hat{\tau}) = 0.224$. It is identical to the expected Fisher information

(see Lemma 1 in [Dandl and Hothorn 2025](#)) evaluated at $\hat{\tau}$. The resulting 95% Wald interval obtained is $(-0.390, 0.486)$. Because the confidence interval is not fully contained in $(-0.36, 0.36)$, the unadjusted analysis cannot reject H_0 .

2.2. Nonparanormal adjusted marginal inference

Adjusting for weight at baseline (variable x) in nonparanormal adjusted marginal inference (with transformation function in Bernstein form of order six) is conducted by fitting a marginal model `m1` for baseline weight in addition to the marginal model for the outcome `m0` and estimating all marginal and Copula parameters simultaneously with `Mmlt()`:

```
R> m1 <- BoxCox(x ~ 1, data = immun)
R> m <- Mmlt(m0, m1, formula = ~ 1, data = immun)

R> (cf1 <- coef(m)["y.w100"])          ### marginal adjusted Cohen's d

      y.w100
-0.005104

R> sqrt(diag(vcov(m))["y.w100"])      ### observed

      y.w100
0.1757

We obtain  $\hat{\tau} = -0.005$  as marginal adjusted Cohen's  $d$ . The standard error based on the observed Fisher information is 0.176.

Given an estimate of the copula parameters in  $\Lambda$  (see Sec. 2.2 in Dandl and Hothorn 2025), we can confirm that the observed Fisher information is equal to the expected Fisher information evaluated at the maximum-likelihood estimates  $\hat{\tau}$  and  $\hat{\lambda}$ :

R> lambda <- c(unclass(coef(m, type = "Lambdapar")))
R> sqrt(2/nrow(immun) * ((1 + lambda^2)*cf1^2 + 8)/(4*(1 + lambda^2))) ### expected

      y.w100
0.1754
```

The Wald interval $(-0.349, 0.339)$ is completely contained in the equivalence interval and thus the absence of an effect of Chloramine on weight can be inferred:

```
R> confint(m)["y.w100",]

      2.5 %   97.5 %
-0.3494  0.3392
```

This finding is in accordance with previous research by [Guo et al. \(2011\)](#). The reduction in standard error comes from the high correlation between the outcome Y and the covariate which can be estimated via

```
R> coef(m, type = "Corr")
```

```
, , 1
```

```
      y      x
y 1.00 0.62
x 0.62 1.00
```

The coefficient of determination R^2 (see Section 2.2 in [Dandl and Hothorn 2025](#)) obtained by

```
R> Omega <- as.array(coef(m, type = "Lambda"))[, , 1]
R> 1 - Omega[nrow(Omega), ncol(Omega)]^(-2)

[1] 0.3844
```

suggests an improvement of the conditional over the marginal model.

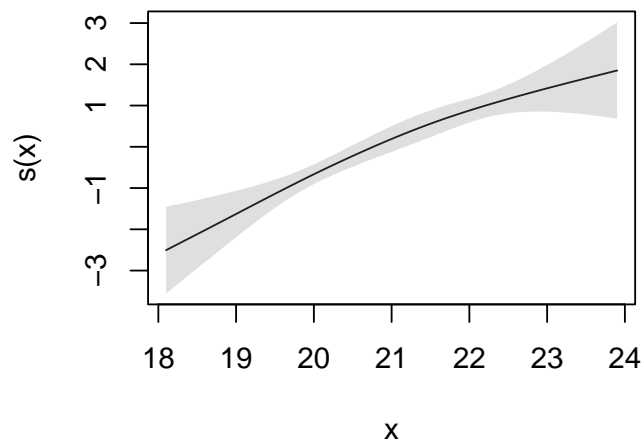
2.3. Model diagnosis

Model diagnosis can be conducted by fitting an additive transformation model for Y using the **tramME** package:

```
R> mad <- BoxCoxME(y ~ w + s(x), data = immun)
```

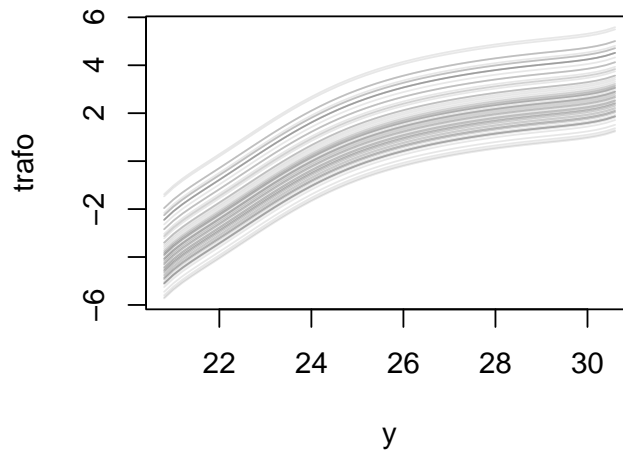
Inspecting the monotonicity of the estimated smooth functions of the baseline covariate, allows to assess the copula fit. The following plot shows the smooth function for baseline weight estimated from the additive transformation model:

```
R> plot(smooth_terms(mad))
```



Because the functions are monotonically increasing, the copula fit seems to be appropriate. The marginal model `m0` for Y relies on the normality assumption and assumes that a linear transformation is sufficient. Via `plot()`, linearity of the transformation function of Y can be assessed using the additive transformation model:

```
R> plot(mad)
```



The plot suggests that a more flexible model allowing a nonlinear transformation h for Y might be better suited. This can be done in the **tram** package using `BoxCox()` for fitting a marginal model. The function estimates a nonlinear transformation function for Y in the form of polynomials in Bernstein form of order M ($M = 6$ is the default):

```
R> m0 <- BoxCox(y ~ w, data = immun)
R> m <- Mmlt(m0, m1, formula = ~ 1, data = immun)
```

The coefficient of determination R^2 can be obtained via

```
R> Omega <- as.array(coef(m, type = "Lambda"))[, , 1]
R> 1 - Omega[nrow(Omega), ncol(Omega)]^(-2)
```

```
[1] 0.5141
```

suggesting an improved model fit over nonparanormal adjusted marginal inference with a linear transformation. This generalised version of Cohen's d is then

```
R> (cf1 <- coef(m)["y.w100"])
```

```
      y.w100
-0.0001844
```

```
R> confint(m) ["y.w100",]
```

```
      2.5 %   97.5 %  
-0.3086  0.3082
```

The obtained treatment effect $\hat{\tau} = -0.000$ with Wald interval $(-0.309, 0.308)$ is interpreted as the mean difference on the latent normal scale in this model. Transforming this estimate to the probabilistic index via $\Phi(\hat{\tau}/\sqrt{2})$, provides a more intuitive interpretation in terms of the probability of obtaining a lower outcome in the control than in the treatment group:

```
R> pnorm(cf1 / sqrt(2))
```

```
y.w100  
0.4999
```

3. Binary outcome: Efficacy study of new chemotherapy

Rödel *et al.* (2012) conducted an efficacy study of combining standard care with a new therapy for patients with rectal cancer. Data from the completed trial (Rödel *et al.* 2015) are available in the **TH.data** package.

623 patients received the Fluorouracil-based standard of care ($W = 0$), and 613 patients received a combination therapy adding Oxaliplatin ($W = 1$). Early effects were assessed based on a binary outcome `ypT0ypN0` – the absence of viable tumour cells in the primary tumour and lymph nodes after surgery:

```
R> CA0surv$ypT0ypN0 <- factor(CA0surv$path_stad == "ypT0ypN0")
```

3.1. Unadjusted marginal inference

Rödel *et al.* (2012) reported an unadjusted odds ratio of 1.4 with 95% confidence interval (1.02, 1.92) based on a Cochran-Mantel-Haenszel χ^2 test (without continuity correction) stratified for lymph node involvement (positive vs. negative) and clinical T category (1–3 vs. 4). Similar results can be obtained by fitting a binary logistic regression model:

```
R> mg_w <- glm(ypT0ypN0 ~ randarm,  
+             data = CA0surv, family = binomial())  
R> exp(coef(mg_w) ["randarm5-FU + Oxaliplatin"])
```

```
randarm5-FU + Oxaliplatin  
          1.422
```

```
R> exp(confint(glht(mg_w), calpha = univariate_calpha())$confint[2,-1])
```

```
      lwr   upr  
1.037 1.949
```

An equivalent model can be obtained with `Polr()` using the **tram** package. The outcome variable `ypT0ypN0` contains 48 missings. For `Polr()`, the missing values in the outcome are retained in the data set, but are ignored in the log-likelihood by choosing `na.action = na.pass`:

```
R> mpCR <- Polr(ypT0ypN0 ~ randarm, data = CA0surv, na.action = na.pass,
+             method = "logistic")
R> exp(coef(mpCR)["randarm5-FU + Oxaliplatin"])

randarm5-FU + Oxaliplatin
               1.422

R> exp(confint(glht(mpCR, coef. = function(...) coef(..., fixed = FALSE)),
+             calpha = univariate_calpha())$confint)

               Estimate   lwr   upr
randarm5-FU + Oxaliplatin    1.422 1.037 1.949
attr(,"conf.level")
[1] 0.95
attr(,"calpha")
[1] 1.96
```

3.2. Nonparanormal adjusted marginal inference

Information on six potentially prognostic covariates is available: `age`, `sex` (`geschlecht`), ECOG performance status (`ecog_o`), distance to the anal verge of the tumour (`bentf`) and the two stratum variables lymph node involvement (`strat_n`) and clinical T category (`strat_t`). For nonparanormal adjusted marginal inference, we define one marginal model for each covariate: a Box-Cox-type model for continuous variable `age`, and binary or ordinal probit models for the remaining binary or ordinal covariates. The ECOG performance status and the distance to the anal verge have 14 missing values, respectively. As for the outcome model, all missing values are retained in the data set but ignored in the marginal log-likelihoods (`na.action = na.pass`):

```
R> mage <- BoxCox(age ~ 1, data = CA0surv)
R> msex <- Polr(geschlecht ~ 1, data = CA0surv, method = "probit")
R> CA0surv$ecog_o <- as.ordered(CA0surv$ecog_b)
R> mecog <- Polr(ecog_o ~ 1, data = CA0surv, na.action = na.pass,
+             method = "probit")
R> mentf <- Polr(bentf ~ 1, data = CA0surv, na.action = na.pass,
+             method = "probit")
R> mT <- Polr(strat_t ~ 1, data = CA0surv, method = "probit")
R> mN <- Polr(strat_n ~ 1, data = CA0surv, method = "probit")
```

The joint model is fitted with `Mmlt`. It is important to note that, unlike the marginal models, this is not a complete cases analysis. If, for example, the outcome is missing for one observation, the corresponding datum is still taken into account. The log-likelihood contribution is

obtained by integrating out the missing dimension from the model. Details on the estimation of the mixed continuous-discrete likelihood are given in [Hothorn \(2024\)](#).

Here, we fit the model

```
R> ### results in the paper were produced using M = 250
R> ### to reduce CRAN checking times, we use M = 50 here
R> ### we also speed up things by allowing the optimiser to
R> ### stop early
R> fastoptH <- mltoptim(abstol = 1e-3, reltol = 1e-3, hessian = TRUE)
R> m <- Mmlt(mage, msex, mecog, mentf, mT, mN, mpCR,
+           data = CAOsurv, args = list(type = "ghalton", M = 50),
+           optim = fastoptH)
R> prm <- "ypT0ypN0.randarm5-FU + Oxaliplatin"
R> exp(coef(m)[prm])
```

```
ypT0ypN0.randarm5-FU + Oxaliplatin
1.402
```

and compute an adjusted confidence interval for the marginal log-odds ratio

```
R> ci <- confint(glht(m, coef. = function(...) coef(..., fixed = FALSE)),
+               calpha = univariate_calpha())$confint
R> exp(ci[prm, -1])

      lwr      upr
1.023 1.922
```

The corresponding odds ratio 1.402 with Wald interval (1.023, 1.922) is very close to the initial results obtained by [Rödel *et al.* \(2012\)](#). The reason is that none of the six variables (including the stratum variables) carries strong prognostic information. The covariate with the largest association with the binary outcome can be identified by ranking the covariates according to their linear correlation after transformation to normality:

```
R> mr <- as.array(coef(m, type = "Cor"))["ypT0ypN0", , 1]
R> i <- which.max(abs(mr[-length(mr)]))
R> (ni <- names(mr)[i])
```

```
[1] "ecog_o"
```

```
R> (mr <- mr[i])
```

```
ecog_o
-0.1521
```

ecog_o is the variable with the largest association with the outcome but its linear correlation under transformation to normality only has a value of $\hat{\rho}_{J, \text{ECOG}} = -0.152$. Consequently, adjusting for covariates did not improve fit nor precision, reflected in a low R^2 :


```
R> Omega <- as.array(coef(m, type = "Lambda"))[, , 1]
R> 1 - Omega[nrow(Omega), ncol(Omega)]^(-2)
```

```
[1] 0.03043
```

However, the standard error did not increase despite the adjustment for multiple covariates, as the following comparison to the standard error of the marginal model shows:

```
R> sqrt(vcov(mpCR))
```

```

                                randarm5-FU + Oxaliplatin
randarm5-FU + Oxaliplatin                                0.1611
```

```
R> sqrt(vcov(m)[prm, prm])
```

```
[1] 0.1609
```

4. Survival outcome: Longevity study of male fruit flies

Partridge and Farquhar (1981) conducted a study on the sexual behavior of fruit flies. The aim was to investigate whether increased sexual activity leads to shorter life spans for male fruit flies. They randomly assigned a total of 125 male fruit flies into five groups of 25 flies: males that live alone, males that live with one or eight receptive females, and males that live with one or eight nonreceptive females. The data set `FruitFlies` is available in the **Stat2Data** package:

```
R> data("FruitFlies", package = "Stat2Data")
```

We follow Negassa and Hanley (2007) and compare only the two groups with eight female flies added that were either all nonreceptive ($W = 0$) or receptive ($W = 1$):

```
R> flies <- FruitFlies
R> flies <- flies[flies$Treatment %in% c("8 virgin", "8 pregnant"),]
R> flies$Treatment <- flies$Treatment[, drop = TRUE]
R> flies$Longevity <- as.double(flies$Longevity)
R> flies$survival <- Surv(flies$Longevity)
```

Of interest is the survival time Y of male flies in days. Because thorax length of the male flies is strongly associated with longevity, it is used as covariate in this analysis.

4.1. Unadjusted marginal inference

The following code fits a marginal Cox proportional hazards model for time to death, with baseline log-cumulative hazard function in Bernstein form of order six:

```
R> coxph_w <- Coxph(survival ~ Treatment, data = flies)
R> coef(coxph_w)          ### log-hazard ratio
```

```
Treatment8 virgin
      2.087
```

```
R> confint(coxph_w)      ### Wald
```

```
      2.5 % 97.5 %
Treatment8 virgin 1.297  2.877
```

The marginal treatment effect is defined as a log-hazard ratio $\hat{\tau} = 2.087$ with 95% Wald interval (1.297, 2.877). Consequently, the hazard of dying in the sexually active group is around $\exp(\hat{\tau}) = 8.060$ times higher than for the nonactive group.

4.2. Nonparanormal adjusted marginal inference

For nonparanormal adjusted marginal inference, we parameterise the transformation function h_1 in the marginal model of thorax length in Bernstein form with order six. The joint distribution of both variables was expressed by a Gaussian copula model:

```
R> xmod <- BoxCox(Thorax ~ 1, data = flies)
R> m <- Mmlt(xmod, coxph_w, data = flies, formula = ~ 1)
R> (cf1 <- coef(m)["survival.Treatment8 virgin"])
```

```
survival.Treatment8 virgin
      1.964
```

```
R> (ci1 <- confint(m)["survival.Treatment8 virgin",])
```

```
      2.5 % 97.5 %
1.384  2.544
```

The log-hazard ratio of nonparanormal adjusted marginal inference is 1.964, with a shorter 95% Wald interval of (1.384, 2.544). The coefficient of determination indicates that thorax length is highly prognostic:

```
R> Omega <- as.array(coef(m, type = "Lambda"))[, , 1]
R> 1 - Omega[nrow(Omega), ncol(Omega)]^(-2)

[1] 0.678
```

4.3. Model diagnosis

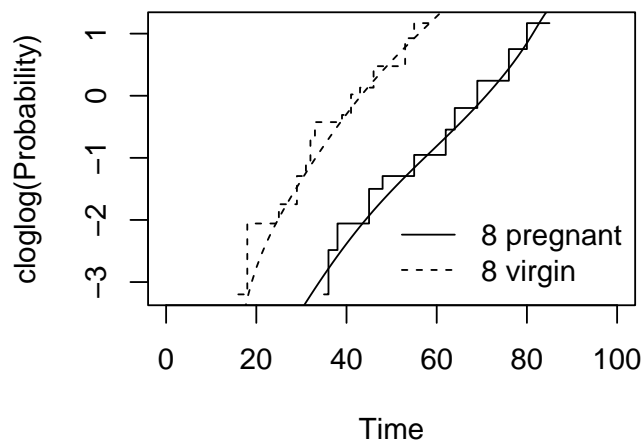
Model diagnosis is conducted on the level of the marginal model for the outcome and on the level of the Gaussian copula parameterising the joint model.

On the level of the marginal model, we compare model-based and nonparametric estimators of the distribution functions of time-to-death on the complementary log-log scale. Plots of these functions are parallel under proportional hazards:

```

R> q <- 0:100
R> cols <- c("grey20", "grey70")
R> ### nonparametric
R> plot(q, log(-log(1 - ecdf(subset(flies, Treatment == "8 pregnant")$survival[,1])(q))),
+       main = "", xlab = "Time", type = "S", lwd = 1,
+       ylab = "cloglog(Probability)")
R> lines(q, log(-log(1 - ecdf(subset(flies, Treatment == "8 virgin")$survival[,1])(q))),
+       type = "S", lty = 2, lwd = 1)
R> legend("bottomright", lty = c(1, 2),
+       legend = levels(flies$Treatment), bty = "n")
R>
R> ### model-based
R> nd <- expand.grid(survival = q, Treatment = sort(unique(flies$Treatment)))
R> nd$h <- predict(as.mlt(coxph_w), newdata = nd, type = "trafo")
R> fm <- nd$Treatment == "8 virgin"
R> lines(nd$survival[fm], nd$h[fm], lty = 2)
R> fm <- nd$Treatment == "8 pregnant"
R> lines(nd$survival[fm], nd$h[fm], lty = 1)

```



From the obtained plot, we see that the smooth model-based transformation functions closely follow the transformed ECDFs (step function). This reflects that the transformation function fits well. The ECDF-based step functions are parallel, revealing that the proportional hazards function of the Cox model is appropriate.

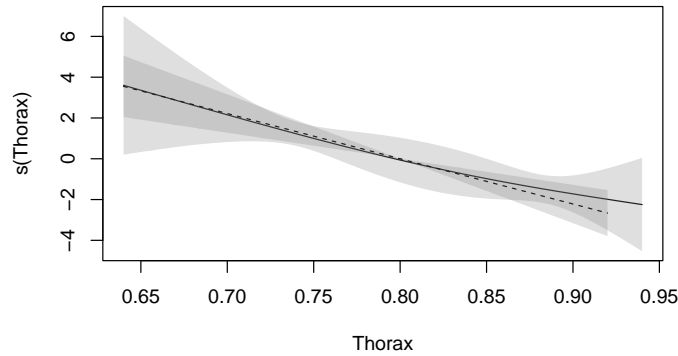
To assess the fit of the Gaussian copula, we fit a more flexible additive transformation model for the outcome given the covariate thorax length as splines for each treatment group:

```

R> m <- CoxphME(survival ~ s(Thorax, k = 5), data = flies,
+               subset = Treatment == levels(Treatment)[1])
R> plot(smooth_terms(m))

```

```
R> m <- CoxphME(survival ~ s(Thorax, k = 5), data = flies,
+               subset = Treatment == levels(Treatment)[2])
R> plot(smooth_terms(m), add = TRUE, lty = 2)
```



Because the smooth functions reflecting the effect of thorax length in both treatment group are monotone, the Gaussian copula seem to be appropriate.

5. Conclusion

The present vignette complements the theoretical work in [Dandl and Hothorn \(2025\)](#) by discussing three application cases of nonparanormal adjusted marginal inference with diverse outcomes. It was showed that nonparanormal adjusted marginal inference can identify the marginal effect of the treatment of interest with higher precision (narrower confidence intervals) as compared to unadjusted analysis if at least one relevant prognostic covariate exists. In addition, model diagnosis tools were discussed that equip users with helpful tools to assess model fits and to examine underlying model assumptions.

References

- Dandl S, Hothorn T (2025). “Nonparanormal Adjusted Marginal Inference.” *Technical report*, arXiv 2503.01657. [doi:10.48550/arXiv.2503.01657](#).
- Guo TL, Germolec DR, Collins BJ, Luebke RW, Auttachoat W, Smith MJ, White KL (2011). “Immunotoxicological Profile of Chloramine in Female B6C3F1 Mice When Administered in the Drinking Water for 28 Days.” *Journal of Immunotoxicology*, 8(4), 381–388. [doi:10.3109/1547691x.2011.622317](#).
- Hothorn T (2024). “On Nonparanormal Likelihoods.” *Technical report*, arXiv 2408.17346. [doi:10.48550/arXiv.2408.17346](#).
- National Toxicology Program (2000). “28-Day Study of Chloramine (10599-90-3) in MICE:B6C3F1.” Last visited 2024-12-22, URL https://tools.niehs.nih.gov/cebs3/views/index.cfm?action=main.dataReview&bin_id=1179.

- Negassa A, Hanley JA (2007). “The Effect of Omitted Covariates on Confidence Interval and Study Power in Binary Outcome Analysis: A Simulation Study.” *Contemporary Clinical Trials*, **28**(3), 242–248. doi:[10.1016/j.cct.2006.08.007](https://doi.org/10.1016/j.cct.2006.08.007).
- Partridge L, Farquhar M (1981). “Sexual Activity Reduces Lifespan of Male Fruitflies.” *Nature*, **294**(5841), 580–582. doi:[10.1038/294580a0](https://doi.org/10.1038/294580a0).
- Rödel C, Graeven U, Fietkau R, Hohenberger W, Hothorn T, Arnold D, Hofheinz RD, Ghadimi M, Wolff HA, Lang-Welzenbach M, Raab HR, Wittekind C, Ströbel P, Staib L, Wilhelm M, Grabenbauer GG, Hoffmanns H, Lindemann F, Schlenska-Lange A, Folprecht G, Sauer R, Torsten Liersch on behalf of the German Rectal Cancer Study Group (2015). “Oxaliplatin Added to Fluorouracil-based Preoperative Chemoradiotherapy and Postoperative Chemotherapy of Locally Advanced Rectal Cancer (the German CAO/ARO/AIO-04 study): Final Results of the Multicentre, Open-label, Randomised, Phase 3 Trial.” *The Lancet Oncology*, **16**(8), 979–989. doi:[10.1016/S1470-2045\(15\)00159-X](https://doi.org/10.1016/S1470-2045(15)00159-X).
- Rödel C, Liersch T, Becker H, Fietkau R, Hohenberger W, Hothorn T, Graeven U, Arnold D, Lang-Welzenbach M, Raab HR, Sülberg H, Wittekind C, Potapov S, Staib L, Hess C, Weigang-Köhler K, Grabenbauer GG, Hoffmanns H, Lindemann F, Schlenska-Lange A, Folprecht G, Sauer R (2012). “Preoperative Chemoradiotherapy and Postoperative Chemotherapy With Fluorouracil and Oxaliplatin Versus Fluorouracil Alone in Locally Advanced Rectal Cancer: Initial Results of the German CAO/ARO/AIO-04 Randomised Phase 3 Trial.” *The Lancet Oncology*, **13**(7), 679–687. doi:[10.1016/s1470-2045\(12\)70187-0](https://doi.org/10.1016/s1470-2045(12)70187-0).
- Wellek S (2010). *Testing Statistical Hypotheses of Equivalence and Noninferiority*. 2nd edition. Chapman and Hall/CRC, Boca Raton, Florida, U.S.A. doi:[10.1201/ebk1439808184](https://doi.org/10.1201/ebk1439808184).

Affiliation:

Susanne Dandl & Torsten Hothorn
Institut für Epidemiologie, Biostatistik und Prävention
Universität Zürich
Hirschengraben 84, CH-8001 Zürich, Switzerland
Torsten.Hothorn@R-project.org