Weighted expectile regression for right-censored data

Alexander Seipp1 | Verena Uslar2 | Dirk Weyhe2 | Antje Timmer1 | Fabian Otto-Sobotka1

1Division of Epidemiology and Biometry, Faculty of Medicine and Health Sciences, Carl von Ossietzky University Oldenburg, Oldenburg, Germany
2University Hospital for General and Visceral Surgery, Pius-Hospital Oldenburg, Oldenburg, Germany

Correspondence
Fabian Otto-Sobotka, Division of Epidemiology and Biometry, Faculty of Medicine and Health Sciences, Carl von Ossietzky University Oldenburg, Ammerländer Heerstr. 114-118, 26129 Oldenburg, Germany.
Email: fabian.otto-sobotka@uni-oldenburg.de

Funding information
Deutsche Forschungsgemeinschaft, Grant/Award Number: SO1313/1-1

Expectile regression can be used to analyze the entire conditional distribution of a response, omitting all distributional assumptions. Among its benefits are computational simplicity, efficiency, and the possibility to incorporate a semiparametric predictor. Due to its advantages in full dataset settings, we propose an extension to right-censored data situations, where conventional methods typically focus only on mean effects. We propose to extend expectile regression with inverse probability weights. Estimates are easy to implement and computationally simple. Expectiles can be converted to more easily interpreted tail expectations, that is, the expected residual life. It provides a meaningful effect measure, similar to the hazard rate. The results from an extensive simulation study are presented, evaluating consistency and sensitivity to violations of assumptions. We use the proposed method to analyze survival times of colorectal cancer patients from a regional certified high volume cancer center.

KEYWORDS
colorectal cancer, inverse probability of censoring, inverse probability weights, iteratively weighted least squares, P-splines

1 | INTRODUCTION

Distributional regression models can be used to model covariate effects beyond average effects, as calculated by the classical mean regression. Generalized Additive Models for Location Scale and Shape (GAMLSS),1 quantile regression2 and expectile regression3 are three main representatives of distributional regression. GAMLSS is a parametric regression approach in which every parameter of a prespecified conditional distribution of the response is modeled. In contrast, quantile regression nonparametrically estimates conditional quantiles. A set of estimated quantiles provides a more complete picture of the distribution and its changes than least squares regression. Covariate effects on lower and upper conditional quantiles describe tail behavior of the distribution.

Expectile regression is an emerging alternative to quantile regression. While quantile regression generalizes median regression, expectile regression is a generalization of mean regression. It inherits the general attributes and modeling flexibility of mean regression, but a lot of the strict assumptions of mean regression are removed. We can estimate expectile regression models with a weighted least-squares method which makes the estimates usually more efficient than regression quantiles. The structure also allows for the straightforward implementation of semiparametric predictors which offer a very flexible construction of the model by possibly including P-splines or Markov random fields as well as parametric effects.4 The latter is much more challenging in a quantile regression setting, at least for frequentist estimates. However, expectile regression offers the same advances as quantile regression in that we can model the entire conditional distribution.

This is an open access article under the terms of the Creative Commons Attribution-NonCommercial License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited and is not used for commercial purposes.

© 2021 The Authors. Statistics in Medicine published by John Wiley & Sons Ltd.
distribution of a response variable with a set of expectiles. With expectile regression, we can estimate the heteroscedasticity that stays hidden in mean regression or we can focus specifically on the tail behavior of the response. The main advantage of quantile regression, on the other hand, is the inherent robustness that any least-squares method will lack.

Compared with quantiles or the GAMLSS model, one disadvantage of single expectiles is their interpretation. However, a dense set of expectiles can be used to estimate the distribution function. Expectiles can also be used to efficiently calculate tail means of a distribution. The upper tail mean will be denoted as expected residual life (also known as the mean residual life, remaining life expectancy). It has various applications in medical, demographic, actuarial and reliability research.

As quantile and expectile regression are nonparametric methods, the estimates for different asymmetry levels will usually be computed separately. While there should be monotonicity in the estimates, in practice neighboring quantile curves often cross. This can also be observed to a lesser extent in expectile regression. In response, several alternative procedures for expectile regression have been developed that enforce noncrossing estimates by either estimating a location-scale model or estimating all expectiles simultaneously with monotonicity constraints.

As in the full data setting, distributional regression could add substantial value to survival analysis if properly extended to handle censored observations. There are already some promising approaches to extend quantile regression. Portnoy developed a quantile regression analog to the Kaplan-Meier estimator. Peng and Huang proposed an analog to the Nelson Aalen estimator instead.

Survival analysis deals with specific missing data problems. A very general approach to handle missing data is the use of inverse probability weights. In this approach, we weight each complete observation with the inverse of the probability that it is not missing. Inverse probability weights can be used in combination with M-estimators. See Tsiatis for extensive theoretical foundations. In the context of right-censoring, this method is called inverse probability of censoring (IPC). IPC dates back to Koul et al, who used IPC to propose an Accelerated Failure Time (AFT) model without distributional assumptions by weighting the responses. Subsequent authors weighted the corresponding loss instead. Zhou applied IPC weights to least squares regression and derived asymptotic properties. Huang et al investigated the use of IPC weights for median regression. In this article, we investigate an extension of expectile regression with IPC weights to handle right-censored time-to-event data.

We demonstrate the benefits of our method on a data set of colon cancer patients. Colon cancer and rectal cancer are referred to as colorectal cancer. The medical standards in Germany for colorectal cancer patients are guided by the S3 guidelines of the Association of the Scientific Medical Societies (AWMF). It includes surgical resection and neoadjuvant and adjuvant therapies. The most effective therapy for colorectal cancer is the surgical resection of the affected part of the intestine with radical regional excision of adjacent lymph node stations. Neoadjuvant radiochemotherapy is carried out preoperatively. Depending on the location and stage of colorectal cancer, neoadjuvant therapy shows clear benefits for disease-free and overall survival. Neoadjuvant therapies are currently only recommended for rectal cancer in the middle and lower third of the rectum. All other primary carcinomas of the colon are treated adjuvantly after the operation. The guideline reports clear advantages of adjuvant chemotherapy for patients with stage III colon cancer, while the data are more unclear at stage II.

Insights like these are typically generated from clinical studies with conventional analysis methods like Kaplan-Meier estimation, proportional hazards models, or AFT models. As an example, Brandi et al conducted a literature review with meta-analysis and concluded that adjuvant therapy for colorectal cancer improves overall survival. This conclusion is based on a meta-analysis of three randomized trials: Nordlinger et al, who used Kaplan-Meier estimates for primary analysis and Mitry et al, who did a pooled analysis of two studies with a Cox proportional hazards model. The use of both methods is common practice and both have their advantages, but also certain disadvantages. The Kaplan-Meier estimator cannot be used if the responses are influenced by multiple covariates and stratification is impractical. The Cox model assumes proportionality of hazard functions between groups. This assumption can often not be verified in practice. It is unclear from reading the article whether Mitry et al checked the proportional hazards assumption or not. Another standard method of analysis is the AFT model, which parametrically models the event time on the log scale. Covariate effects are assumed to be location-shifts, which is the case if the residuals are distributed identically. If variance, skewness, or other characteristics of the distribution change over time, the AFT model provides inadequate estimates to describe more than the covariate effect on the center.

The remainder of this article is structured as follows: First, we introduce expectile regression. Then we present our estimator for right-censored data. We give a summary of M-quantiles and asymptotic results that exist in the literature. Section 3 contains the simulation results for our estimator. We then apply our method to colon cancer data collected at our hospital. The article is concluded with a discussion of our findings.
2 | METHOD DEVELOPMENT

2.1 | Expectile regression

In classic linear regression we minimize the sum of squared residuals,

$$\hat{\beta} = \arg\min_\beta \sum_{i=1}^n (t_i - x_i^T \beta)^2,$$

where $t_1, \ldots, t_n$ are the observed responses and the corresponding random variables are written in capital letters. The parameter vector is $\beta \in \mathbb{R}^p$, $x_i \in \mathbb{R}^p$ denotes the $i$th row of the design matrix $X \in \mathbb{R}^{n \times p}$.

It should be emphasized that $x_i^T \beta$ is the conditional mean of the response $T_i$ given some covariate vector $x_i$. The estimate $\hat{\beta}$ therefore also refers to changes in the conditional mean. We denote this type of regression by “mean regression.”

If the residuals are not identically distributed, the mean might not be the only parameter of interest. When we summarize information of univariate data, we calculate multiple summary statistics, including the mean, median, other quantiles and skewness. Just as in the univariate case, additional information about the whole conditional distribution is useful to represent the data. Quantile regression can be used to estimate conditional quantiles. In quantile regression, we do not assume a common parametric distribution. It is therefore suitable for use if residuals are not identically distributed (e.g., heteroscedastic). Another alternative approach is expectile regression. Expectiles are a generalization of the expectation, depending on an asymmetry parameter $\tau \in [0, 1]$. In addition to the center ($\tau = 0.5$), they describe the lower tail ($\tau < 0.5$) and the upper tail ($\tau > 0.5$) of the distribution. As with quantiles, expectiles characterize a distribution and we can estimate the empirical distribution function from a large set of empirical expectiles.

In a regression setting there is a different parameter vector $\beta_\tau$ for each $\tau$. We can estimate them separately with minimization of weighted squared residuals. The weights $w_i$ are asymmetric and have a different penalty for positive and negative residuals:

$$\hat{\beta}_\tau = \arg\min_\beta \sum_{i=1}^n w_i(t_i - x_i^T \beta_\tau)^2,$$

$$w_i = \begin{cases} \tau & t_i \geq x_i^T \beta_\tau \\ 1 - \tau & t_i < x_i^T \beta_\tau \end{cases}.$$

Let $\mu(\tau, T, x)$ be the $\tau$-expectile of $T_i$ given covariates $x$. If the model is correctly specified, we have $\mu(\tau, T, x) = x^T \beta_\tau$ and $\partial \mu(\tau, T, x)/\partial x = \beta_\tau^T$. Therefore, a set of $\beta_\tau$ shows changes in different parts of the conditional distribution. In contrast to mean or quantile regression, however, the interpretation of $\beta_\tau$ is not as intuitive. The expected residual life, a more meaningful effect measure, is described in Section 2.4.

After differentiation with respect to $\beta_\tau$, we get $\hat{\beta}_\tau$ as the solution to

$$\sum_{i=1}^n w_i(t_i - x_i^T \hat{\beta}_\tau)x_i = 0,$$

which we can rewrite as

$$\hat{\beta}_\tau = (X^T W_{\tau}(\hat{\beta}_\tau)X)^{-1}X^T W_{\tau}(\hat{\beta}_\tau)t$$

with $W_{\tau}(\hat{\beta}_\tau) = \text{diag}(w_1, \ldots, w_n)$. Since the weights depend on the estimate, we have to solve iteratively for $\hat{\beta}_\tau$. This is easily done by using existing weighted least squares implementations. This is in contrast to quantile regression, which requires a numerical optimization routine.

So far we have only introduced multiple linear expectile regression models. However, the effect of metric covariates is not always linear. Further, a covariate may also hold spatial information. However, the shape of a nonlinear or spatial effect is usually unknown, especially in an additive model. Hence, we opt to approximate the effects of noncategorical covariates with a base and penalty approach. For metric covariates, for example, we define P-spline bases as made popular by Eilers and Marx. The basis matrix $B$ is defined by a dense set of local polynomial splines on an equally spaced grid.
covering the support of a covariate. The nonlinear effect is then estimated by a linear combination of the basis elements and their regression coefficients \( \gamma \). We add a penalty term \( K \) to the estimation to avoid overfitting. For P-splines the penalty consists of the sum of squared second-order differences of neighboring coefficients, that is,

\[
\lambda \gamma^T D^T D \gamma = \lambda \sum_j (\gamma_j - 2\gamma_{j-1} + \gamma_{j-2})^2,
\]

where \( K = D^T D \). \( \lambda \) is a smoothing parameter that controls the balance between fit and penalty. An additive predictor can then be defined as \( \eta = Z \theta \) with design matrix \( Z = (X, B_1, \ldots, B_r) \) and parameter vector \( \theta = (\beta, \gamma_1, \ldots, \gamma_r) \) where \( X \) comprises all parametric effects and an intercept. Given a fixed vector of smoothing parameters, the minimization problem is

\[
\hat{\theta}_r = \arg \min_{\theta_r} \sum_{i=1}^n w_i (t_i - z_i^T \theta_r)^2 + \theta^T P(\lambda) \theta
\]

with \( P(\lambda) = \text{diag}(1, \lambda_1, \ldots, \lambda_r) D^T D \). \( z_i \) is the \( i \)th row of \( Z \). After differentiation with respect to \( \theta \), we get \( \hat{\theta}_r \) as the solution to

\[
\sum_{i=1}^n (w_i [t_i - z_i^T \theta_r] z_i) + P(\lambda) \theta = 0,
\]

which leads to

\[
\hat{\theta}_r = (Z^T W_r(\hat{\theta}_r) Z + P(\lambda))^{-1} Z^T W_r(\hat{\theta}_r) t.
\]

\( \hat{\theta}_r \) can again be computed iteratively. The optimal choice of smoothing parameters was investigated by Schnabel and Eilers. They proposed an adaptation of the Schall algorithm, where expectile regression for a given \( \lambda \) and recalculation of \( \lambda \) based on the estimated residuals iterates.

### 2.2 Expectile regression for right-censored data

Clinical trials often include right-censored observations where a proportion of the event times \( T_i \) is known exactly, but another proportion is only known to be larger than some censoring time \( C \). Our data then have the structure \((Y_1, x_1, \Delta_1), \ldots, (Y_n, x_n, \Delta_n)\) where \( \Delta_i = I(T_i \leq C_i) \) and \( Y_i = \min(T_i, C_i) \), which is referred to as the follow-up time. We assume independence of \( T_i \) and \( C_i \), conditional on \( x_i \). Since we do not know the event times, but only the follow-up times, we cannot use equation (1) directly. We propose using inverse probability of censoring (IPC) weights to account for right-censoring. Let \( v_i = \delta_i / P(C_i > t_i | z_i) \) denote the IPC weights. Our estimate solves

\[
\sum_{i=1}^n v_i w_i [t_i - z_i^T \theta_r] z_i) + P(\lambda) \theta = 0,
\]

which we can again write as

\[
\hat{\theta}_r = (Z^T V W_r(\hat{\theta}_r) Z + P(\lambda))^{-1} Z^T V W_r(\hat{\theta}_r) t
\]

with \( V = \text{diag}(v_1, \ldots, v_n) \). The IPC estimator uses only uncensored observations, which receive larger weights depending on the probability that they are uncensored. Each uncensored observation thereby also represents several censored observations. The idea is to solve an equation that has the same expectation as the left side of equation (1) and can be used with the observed data. The left sides of (1) and (2) have the same expectation, which can be seen by exploiting \( E(\cdot) = E(E(\cdot | T)) \) and \( E(\Delta / v) = 1 \).

Since in practice \( P(C > t | z) = G(t | z) \) is unknown, we replace it with an estimate \( \hat{G}(t | z) \). Huang et al. proposed IPC weights for median regression with right-censored data. They estimated the IPC weights with \( \hat{\delta} = \delta / \hat{G}(t) \), where
\( \hat{G}(t) = \hat{G}(t | z) \) is the Kaplan-Meier estimate of the survival function of the censoring time. They showed that their proposed estimator of the median regression parameters is consistent and asymptotically normal. Zhou\(^{12} \) used IPC weights for mean regression with right-censored data. The weights of Zhou differ slightly from Huang et al. They are identical to the weights in Huang et al for each observation except the last. For the last observation, the weights are instead

\[
\hat{v}(n) = \frac{1}{\lim_{u \to t_n}} \hat{G}(u).
\]

Zhou argues that with this definition no weight is lost. This is in line with Efron\(^{20} \) and his “redistribute-to-the-right” approach to the product-limit estimator. Zhou’s estimator is also shown to be consistent and normally distributed. From now on we will denote Huang et al’s weights with IPCKM and Zhou’s weights with IPCRR.

We want to emphasize two central assumptions:

1. \( C \) is independent of the covariates.
2. \( G(t | z) > 0 \) for the whole support of \( T \).

Adherence to assumption (2) can depend on the study design. Assumption (2) might be violated for Type 1 censoring with a short observation period. We investigate sensitivity to violations of both assumptions with our simulations.

The largest advantage that we see in IPC weights is their simplicity, both conceptional and computational. Parameter estimates can be computed by small modifications in the routines of expectile regression. In the same way, we can adapt automated smoothing parameter choices (asymmetric cross-validation, Schall algorithm, generalized AIC, etc) and model selection criteria. Spiegel et al\(^{21} \) define several selection criteria, including a generalized AIC. We can for example adapt the generalized AIC with

\[
AIC = n \log \left( \frac{1}{n} \sum_{i=1}^{n} v_i w_i (t_i - \hat{z}_T, \hat{\theta})^2 \right) + 2p,
\]

where \( p \) denote the effective degrees of freedom. They are calculated with \( p = \text{tr}(H_r) = \text{tr}(K_r) \). \( H_r \) is the generalized hat matrix that has the same trace as matrix \( K_r \), defined as

\[
K_r = (Z^T V \bar{W}_r (\hat{\theta}, \hat{\lambda}) Z + P(\lambda))^{-1} Z^T V \bar{W}_r (\hat{\theta}, \hat{\lambda}) Z.
\]

This definition differs from Spiegel et al only by adding IPC weights at the appropriate positions.

### 2.3 Asymptotics

In M-estimation\(^{22} \) an estimating equation is used to estimate the center of the distribution. Depending on the choice of the estimating function \( \psi \), we can achieve more robustness than in mean regression. M-quantile regression uses asymmetric weights to generalize M-estimation in the same way as expectile regression generalizes mean regression. We can infer the asymptotic behavior of the IPC estimator with true weights from M-quantile regression, hence we give a brief overview of it.

The conditional M-quantile given \( x_i \) is defined by \( x_i^T \hat{\beta}_r \). The true parameter vector \( \beta_r \) is given by

\[
E(w \psi(T, x, \beta_r)) = 0
\]

and the estimator \( \hat{\beta}_r \) is the solution to

\[
\sum_{i=1}^{n} w_i \psi(T_i, x_i, \beta_r) = 0.
\]

Bianchi and Salvati\(^{23} \) derived asymptotic properties of \( \hat{\beta}_r \). Under regularity conditions we have

\[
\sqrt{n}(\hat{\beta}_r - \beta_r) \overset{d}{\to} N(0, \Sigma).
\]
where the covariance matrix of $\hat{\beta}_r$ is $\Sigma/n$. Bianci and Salvati suggest estimating the covariance matrix with a sandwich estimator.

First, consider linear expectile regression and assume the true IPC weights to be known. Consistency and asymptotic normality follow directly from this result. We get linear expectile regression with $\psi(T, x, \beta_r) = (T - x^T \beta_r)x$ and the IPC estimator with $\psi(T, x, \hat{\beta}_r) = v(T - x^T \hat{\beta}_r)x$. Note that the true parameters for both expectile regression with full data and IPC weighted regression with right-censored data are identical since their expectation in (3) is the same. This can be shown again by a conditioning argument. Otto-Sobotka et al.\textsuperscript{24} considered semiparametric M-quantile regression and provided an adapted covariance matrix. For IPC-weighted expectile regression their estimator reduces to

$$
\text{Cov}(\hat{\theta}_r) = \frac{n}{n-1} A^{-1} B A^{-1},
$$

$$
A = Z^T VW_r(\hat{\theta}_r) Z + P(\lambda),
$$

$$
B = Z^T (VW_r(\hat{\theta}_r) \text{diag}(T - Z\hat{\theta}_r))^2 \text{diag}(I - H_r)^{-1} Z.
$$

(4)

In our proposition, $v$ is replaced by an estimate $\hat{v}$ and the given asymptotics do not apply. We propose estimation of the covariance matrix with the sandwich estimator given by Otto-Sobotka et al.\textsuperscript{24} where $\hat{v}$ is used as a plug-in estimator instead of $v$. We investigate the performance of this covariance estimator and the consistency of the IPCKM and IPCRR estimator through simulations.

### 2.4 Calculating the expected residual life from expectiles

As described above, the interpretation of single expectiles is considered less intuitive than that of single quantiles. However, we can efficiently calculate the expected shortfall $ES(q)$ from expectiles.\textsuperscript{25} The expected shortfall is a financial risk measure and is defined as

$$
ES(q) = E(T|T < t),
$$

where $q = P(T \leq t)$. Taylor’s results also include the formula to calculate the upper tail version of the expected shortfall, which is related to the expected residual life. The expected residual life is defined by

$$
\text{ERL}(t, x_0) = E(T - t|T > t, x_0).
$$

The relationship to expectiles is given with

$$
\text{ERL}(t, x_0) = \frac{1 - \tau}{(2\tau - 1)(1 - q)} (t - \mu(0.5, T, x_0)),
$$

(5)

where $\tau$ solves $\mu(\tau, T, x_0) = t$ and $q = P(T \leq t|x_0)$ for some fixed covariate vector $x_0$. The expected residual life can be estimated using expectile regression for a dense sequence of $\tau$ values between 0 and 1 and estimating the conditional densities from the predicted expectiles. An approach of estimating a density $f(t)$ from expectiles $\mu(\tau)$ was shown in Schnabel and Eilers\textsuperscript{6} by first considering their relationship through the equation

$$
(1 - \tau) \int_{-\infty}^{\mu(\tau)} (u - \mu(\tau))f(u)\, du + \tau \int_{\mu(\tau)}^{\infty} (u - \mu(\tau))f(u)\, du = 0
$$

and using a discrete approximation in combination with a roughness penalty to estimate the density. The expected residual life can be estimated from the density directly. However, the ERL can be calculated through Equation (5) without necessarily relying on upper expectiles, which will be shown to have estimates with larger bias on average.

In survival analysis, we commonly state hazard rates or hazard ratios to quantify the risk of death. However, doctors and patients might be more interested in the survival time itself. Proportional expected residual life models have been used as an alternative to proportional hazard models, for example, for analysis of overall survival of patients with lung cancer and glioblastoma multiforme.\textsuperscript{26,27} We aim to demonstrate the utility of ERL in Section 4.
TABLE 1 Parameter choices in the four base scenarios

| Base scenario | Hazard          | $\varepsilon$ | $\sigma$ | $\alpha$ |
|---------------|-----------------|---------------|----------|----------|
| 1             | Constant        | Gumbel        | 1        | 2.50     |
| 2             | Concave increasing | Gumbel       | 0.87     | 2.47     |
| 3             | First increasing, then decreasing | Normal | 1.4      | 1.98     |
| 4             | First decreasing, then increasing | Uniform    | 2.8      | 0.77     |

3 | SIMULATION STUDY

We conducted a simulation study with a wide range of scenarios to assess the performance of our proposed estimator.

3.1 | Design

Right-censoring often occurs for strictly positive event times. In that case our estimator should be used on logarithmized event times. Denote $T^*$ as the positive event time in months and $T = \log(T^*)$. For other variables the notation is analogous. We chose four base scenarios and several simulation parameters for modification.

The base scenarios were chosen to correspond to a study lasting roughly 2 years. Approximately 79% of the follow-up times were shorter than one year and approx. 93% were shorter than 2 years. We chose

$$T_i = \alpha + \frac{\alpha}{\log(2)}(X_{i1} - X_{i2}) + \sigma \varepsilon_i,$$

$$C_i = \log(20) + e_i,$$

$$Y_i = \min(T_i, C_i).$$

$T^*_i$ was assumed to follow either an exponential, Weibull, log-normal, or log-uniform distribution. This corresponds to $\varepsilon_i$ being i.i.d. random variables, which follow either a Gumbel distribution, a standard normal distributed or a uniform distribution on $[0, 1]$. $C^*$ was chosen to be exponentially distributed with a mean 20 of months to fit the study duration of 2 years. Accordingly, $e_i$ were i.i.d. Gumbel distributed. The covariates $X_{ij}$ were i.i.d. with

$$X_{ij} \sim U[0, \log(2)], \quad i = 1, \ldots, n, \quad j = 1, 2.$$ 

Therefore, the support for $\exp(X_{ij})$ was $[1, 2]$. We included the factor $\alpha / \log(2)$ so that the covariate effects were standardized and were always between 0 and $\alpha$. $\alpha$ was chosen such that censoring was 40% on average. See Table 1 for an overview of the parameters in the four base scenarios. Note that throughout this article, all numbers shown are rounded to two decimal places.

We chose the distributions for $T^*$ and the values for $\sigma$ such that four different hazard shapes were covered. The hazards were constant (exponential distribution), concave increasing (Weibull distribution), decreasing after a clear maximum (log-normal distribution), and first decreasing, then increasing (log-uniform distribution). Figure 1 shows the hazard functions in the base scenarios. The exponential distribution is included due to its extensive use in survival analysis. Examples for the other shapes can be found, among other fields, for breast cancer. Increasing hazard rates were found for age-specific breast cancer incidence. A hazard shape with a high risk at the beginning that decreases were found after treatment of breast carcinoma. Increasing hazard rates after the initial drop (as seen for the log-uniform distribution) were also found after breast carcinoma treatment, possibly due to cancer relapses. The log-uniform distribution was also chosen since its support is bounded. This results in $G(t|z) > 0.02$ for any $t \in \text{supp}(T)$ and assumption (2) holds. While this assumption also holds in the other scenarios, $G(t|z)$ can be close to 0. Finally, in the base scenarios (and all other scenarios), we chose sample sizes of $n = 200$ and $n = 800$.

Additional scenarios were chosen in which some parameters of the base scenarios were varied, both ceteris paribus and in combination with other variations. In addition to the four base scenarios, we varied five parameters with two levels each, resulting in $4 \cdot 2^5 = 128$ different scenarios. The share of censoring was reduced from 40% to 20% (variation 1), the
FIGURE 1  Hazard functions of $T^*$ for the four base scenarios [Colour figure can be viewed at wileyonlinelibrary.com]

| Variation number | Parameter                  | Base scenario | Variation                                |
|------------------|----------------------------|---------------|------------------------------------------|
| 1                | Censoring                  | 40%           | 20%                                      |
| 2                | Covariate influence on $T$ | Linear        | Fourth degree polynomial                 |
| 3                | Covariate influence on $C$ | Independent   | Linear                                  |
| 4                | Scale                      | $\sigma$      | $0.2\sigma \leq \sigma (X_{i1},X_{i2}) \leq 1.8\sigma$ |
| 5                | Upper bound of censoring   | $\infty$      | log(24)                                  |

covariates’ influence on the event time was nonlinear (variation 2), the covariates were associated with the censoring time (variation 3), the covariates influenced on the scaling parameter (variation 4), and an upper bound of censoring was introduced (variation 5). Table 2 gives an overview of the parameter changes and details are described below.

In variation 1, $\alpha$ was reduced such that the share of censoring was 20% instead of 40% on average. In variation 2, instead of a linear influence of the covariates on $T$, we chose a fourth degree polynomial. We emulated the shape of the effect from real data, which we present in Section 4. It was modeled like the effect of the lymph ratio on the 0.05-expectile of the event time distribution (see Figure 5). We chose

\[
T_i = \alpha + \frac{\alpha}{1.84} (-195.15(X_{i1}^4 - X_{i2}^4) + 279.40(X_{i1}^3 - X_{i2}^3) \\
- 131.32(X_{i1}^2 - X_{i2}^2) + 20.60(X_{i1} - X_{i2})) + \sigma e_i.
\]

We standardized with a constant again, such that the covariate effect was always between 0 and $\alpha$. Variation 3 violated assumption (1) by generating censoring in dependence of the covariates. We opted for

\[
C_i = \log(20) + \frac{\log(20)}{\log(2)}(X_{i1} - X_{i2}) + e_i,
\]

where $e_i$ were i.i.d. Gumbel distributed. As above, the covariate effect was standardized to values between 0 and log(20). Heteroscedasticity of the event time (variation 4) was achieved by replacing $\sigma$ with $\sigma (X_{i1},X_{i2})$, where

\[
\sigma (X_{i1},X_{i2}) = \sigma \left( 1 + \frac{0.8}{\log(2)}(X_{i1} - X_{i2}) \right).
\]

The chosen scale depended linearly on the covariates and varied between $0.2\sigma$ and $1.8\sigma$. We added an upper bound (variation 5) that represented a fixed time limit of 2 years on follow-up. The censoring time was defined by $C_i^u = \min(C_i, \log(24))$. In these scenarios, assumption (2) was violated.
We repeated the simulation 1000 times for each scenario and aggregated the results. For each repetition, full data expectile regression as well as censored data expectile regression with IPCKM and IPCRR weights were conducted. We chose the asymmetry levels 0.01, 0.02, 0.05, 0.1, 0.2, 0.5, 0.8, 0.9, 0.95, 0.98, 0.99. This put the focus on the tails of the distribution, where the most noticeable changes often occur. Predictions for conditional expectiles were made and point estimates and confidence intervals were evaluated. For the point estimates we made a prediction $\hat{\mu}(\tau, T_i, x_i)$, for every observed covariate combination and compared it with the true conditional expectiles. The true expectiles depend on the first partial moment of the distribution, which is known for some distributions. Additional partial moments used as well as the general formula for the calculation of the univariate expectiles can be found in the Appendix. We calculated bias and MSE, defined with

$$\text{Bias}(\tau) = \frac{1}{n} \sum_{i=1}^{n} \hat{\mu}(\tau, T_i, x_i) - \mu(\tau, T_i, x_i),$$

$$\text{MSE}(\tau) = \frac{1}{n} \sum_{i=1}^{n} [\hat{\mu}(\tau, T_i, x_i) - \mu(\tau, T_i, x_i)]^2.$$

We then aggregated bias and MSE for all repetitions. For the confidence intervals we calculated predictions on an equally spaced two-dimensional grid of 100^2 points along the covariates $X_1$ and $X_2$. We constructed 95% confidence intervals at each point on the grid and for each expectile level with

$$\hat{\mu}(\tau, T_i, x_i) \pm u_{0.975}\text{SE}(\hat{\mu}(\tau, T_i, x_i)),$$

where $u_{0.975}$ is the 0.975-quantile of the standard normal distribution and the standard errors were calculated based either on Equation (4) or with a nonparametric bootstrap (200 repetitions). We compared five coverages: IPCKM and IPCRR, both with asymptotic plug-in variance estimator and bootstrapping, as well as classic expectile regression with full data. Our goal was to evaluate the finite sample performance of the plug-in variance estimator in comparison with the nonparametric bootstrap as gold standard. In the results we report the empirical coverage.

In this section we present exemplary results from a few selected scenarios of our simulation study. The entirety of the results are collected in Web Appendix A (Supplementary Materials). We used the software R and the package expectreg, our code is available upon request.

### 3.2 Point estimates

First, we focus on bias and MSE of the point estimates. The results of the base scenarios 1 to 4 are shown in Tables 3, 4 and tables given in Appendix B. Our key findings were the following:

- Both the IPCKM and IPCRR method tended to underestimate regression parameters.
- The IPCRR method performed better than IPCKM.
- Performance improved with increasing sample size.
- Estimates for lower expectiles were the most reliable. Estimates for 98%- and 99%-expectiles were the most unreliable.
- The estimators were sensitive to deviation from assumption (1).
- Bias was introduced if $G(t|z)$ was close to 0.

In all of the scenarios, the bias was negative for most of the distribution. We found a very small positive bias in the extreme lower tails at the 0.01 and 0.02-expectiles. For all expectiles the bias was more negative for the IPC methods in comparison with the artificial estimation using full data. The size of the bias increased the most for the upper tail at the 0.98 and 0.99-expectiles. In the scenarios with a maximum censoring time of log(24) (variation 5) the bias increased in size with a stronger bias again in the extreme upper tail. In scenario 1 the bias for the 0.99-expectile increased from $-0.36$ to $-0.74$ for IPCKM and from $-0.31$ to $-0.57$ for IPCRR. As expected the share of censored observations (variation 1) and the overall sample size had a strong impact on the bias except for the scenarios with a limited censoring time. Overall,
the bias was smaller for the IPCRR method than for the IPCKM estimates. The MSE was comparable in the lower tail for the IPC methods and the full data estimation. In the upper tail, the MSE of the IPCKM and IPCRR was multiples of the MSE using full data. In scenario 4 the MSE was much smaller than for the other scenarios due to the bounded log-uniform event time distribution. A reduction of the share of censoring (variation 1) improved the MSE as expected. In scenario 2, for example, the MSE was reduced by 76% and 78% for both methods, respectively, with a reduced amount of censored observations. The performance of both estimators was also worse in heteroscedastic scenarios (variation 4). In scenario 3, the MSE was particularly high for the 0.99-expectile. It increased to 2.75 (IPCKM) and to 2.04 (IPCRR). For other expectiles it was much lower, for example, 0.31 (IPCKM) and 0.14 (IPCRR) for the mean. As we introduced a nonlinear effect for a covariate which was estimated by P-splines (variation 2), the MSE increased by 0.06 (IPCKM, IPCRR) for the mean and by 0.22 (IPCKM, IPCRR) for the 0.99-expectile. For the increased sample size the reduction in MSE was similar for both methods. In scenario 1 it was reduced between 54% (0.99-expectile) and 72% (0.02-expectile) for IPCKM and between 53% (0.99-expectile) and 74% (0.05-expectile) for IPCRR. In summary, the IPCRR method also outperformed the IPCKM method regarding the MSE.
3.3 Confidence intervals

In Figure 2 some of the results from base scenario 3 with a sample size of 200 are displayed. We show the coverage share for the 0.1, 0.5, and 0.9 expectile along both metric covariates, respectively. The full data confidence intervals did not reach their desired coverage rate on every point of the grid. It ranged from 89% to 100% in total. Averages per expectile level ranged from 91% ($\tau = 0.01$) to 95% ($\tau = 0.5$). The coverage for the censored data methods was worse for the upper expectiles than for lower expectiles. We report a selection of coverages as a mean over the prediction grid. In the case of IPCKM the average coverage was 90% ($\tau = 0.1$), 73% ($\tau = 0.5$), and 52% ($\tau = 0.9$) for plug-in standard errors. For IPCRR it was 95% ($\tau = 0.1$), 90% ($\tau = 0.5$), and 71% ($\tau = 0.9$). When we estimated the standard errors via bootstrapping it resulted in an average coverage of 88% ($\tau = 0.1$), 69% ($\tau = 0.5$), and 47% ($\tau = 0.9$) for IPCKM. For IPCRR it resulted in an average coverage of 93% ($\tau = 0.1$), 87% ($\tau = 0.5$), and 65% ($\tau = 0.9$).

In comparison with scenarios 1 to 3, scenario 4 showed the best coverage. For the 0.1-expectile, average coverage was 95% (full data), 95% (IPCKM, plug-in), 93% (IPCKM, bootstrapping), 95% (IPCRR, plug-in), and 94% (IPCRR, bootstrapping). Average coverage for the 0.9-expectile was 95% (full data), 90% (IPCKM, plug-in), 87% (IPCKM, bootstrapping), 93% (IPCRR, plug-in), and 90% (IPCRR, bootstrapping).

For further results, we refer the reader to Web Appendix A (Supplementary Materials). It contains simulation results for the four base scenarios. The results for the confidence intervals in the additional scenarios are not included, since the bootstrapping was computationally demanding, especially in combination with P-splines and automatic smoothing parameter selection.
APPLICATION: COLON CANCER DATA

We analyzed an anonymized data set with survival times of colon cancer patients. The patients were treated at our specialized cancer center. The patients were registered from September 2009 to March 2019. We excluded 68 cases with incomplete covariate values, resulting in an analysis set with 563 cases.

We focused on two research questions. First, we were interested in the association between chemotherapy and overall survival for different parts of the distribution. Kaplan-Meier analyses for stage III colon cancer suggest that the largest effect could be approximately at the 0.3-quantile. For the exponential distribution, this corresponds to the 0.075-expectile. Expectile regression extends this research question to the multivariate setting.

We were secondly interested in whether the number of examined lymph nodes was associated with the overall survival. An increased number of examined lymph nodes could lead to a more accurate severity assessment of the tumor and better treatment recommendations. If the patient’s lymph nodes are not affected by cancer, the patient typically does not receive chemotherapy.

The lymph nodes were examined at one of two pathology laboratories A and B. While Laboratory B generally examines every lymph node removed during surgery, Laboratory A generally only examines a portion of the available nodes. The different amounts of examined lymph nodes are shown in Figure 3. While the amount of resected lymph nodes varies, removal of at least 12 is recommended. This recommendation for a minimum quantity is normally only undercut if more lymph nodes could not be removed due to surgical-anatomical reasons. There are also upper limits to the number of lymph nodes that can be examined, due to standardized surgical procedures, which depend on the part of the intestine that is resected, and which automatically predetermine the number of resected lymph nodes to a certain extend.

We preselected clinically relevant covariates. Descriptive statistics can be found in Table 5. Patients’ general illness was classified by the scoring system of the American Society of Anesthesiologists (ASA score). An ASA score smaller than 3 is considered a mild general illness, 3 or greater is considered a severe general illness. An anastomotic insufficiency is a postoperative complication that can lower survival chances. The UIC classification of colorectal cancer was used to classify the tumors in stages I to IV. Severity is sorted in ascending order. A patient is classified as stage II if the primary tumor infiltrates surrounding tissue or organs. Patients with metastases in regional lymph nodes are considered at least stage III. Patients with stage IV cancer have at least one remote metastasis. The lymph node ratio is defined as the number of affected lymph nodes divided by the total number of examined lymph nodes. Preexisting cancer refers to independent cancer diseases that have occurred at any point in the patient’s medical history.

Median follow-up was 743 days. Since the follow-up time was short and patients with colon cancer have better survival chances than, for example, lung cancer patients, around two-thirds (66.3%) of the observations were censored. Since the survival time was strictly positive, we used the log follow-up time for the expectile regression models. The log follow-up time is illustrated in Figure 4. Figure 4 shows a large amount of censoring at the end, hence it is likely that parts of the upper tail of the distribution are not represented well. This is not ideal for the IPC estimator, as shown in Section 3. Therefore, we should be careful when interpreting results from the upper tail of the distribution.
TABLE 5  Descriptive statistics of the preselected covariates

|                          | No               | 384 (68.2%) | Yes   | 179 (31.8%) |
|--------------------------|------------------|-------------|-------|-------------|
| Chemotherapy             |                  |             |       |             |
| Yes                      |                  |             |       |             |
| Sex                      | Male             | 242 (43.0%) |       |             |
| Female                   | 321 (57.0%)      |             |       |             |
| ASA classification       | ≤ 2              | 220 (39.1%) |       |             |
| ≥ 3                      | 343 (60.9%)      |             |       |             |
| Anastomotic insufficiency| No               | 535 (95.0%) |       |             |
| Yes                      | 28 (5.0%)        |             |       |             |
| UICC stage               |                  |             |       |             |
| I                        | 107 (19.0%)      |             |       |             |
| II                       | 203 (36.1%)      |             |       |             |
| III                      | 160 (28.4%)      |             |       |             |
| IV                       | 93 (16.5%)       |             |       |             |
| Age                      | Mean             | 70.9        |       |             |
| SD                       | 12.5             |             |       |             |
| Examined lymph nodes     | Mean             | 25.0        |       |             |
| SD                       | 12.0             |             |       |             |
| Lymph node ratio         | Mean             | 0.1         |       |             |
| SD                       | 0.2              |             |       |             |
| Preexisting cancer       | No               | 465 (82.6%) |       |             |
| Yes                      | 98 (17.4%)       |             |       |             |

FIGURE 4  Histograms of log follow-up time (base 10) for patients without event (censored) and with event [Colour figure can be viewed at wileyonlinelibrary.com]
First, we fitted parametric accelerated failure time (AFT) models including every preselected variable (full model). The AFT models were fitted assuming a Weibull, a log-normal, or a log-logistic distribution. Comparing the empirical distribution function of the estimated residuals with the theoretical distribution function, we concluded that the Weibull fit was most reasonable. We then used best subset selection for variable selection. The model with the lowest AIC was chosen. The final parsimonious model excluded sex, anastomotic insufficiency, and preexisting cancer. We also considered an interaction term between chemotherapy and stage III or higher. However, we excluded the interaction term in the selection process. The resulting estimates and corresponding 95% confidence intervals of the Weibull model can be found in Table 6. Reference categories correspond to the first entries in Table 5. In the parsimonious model, longer survival times (log days) were associated with receiving chemotherapy (estimate: 0.91, SE: 0.22). The association with the number of examined lymph nodes was also positive (estimate: 0.03, SE: 0.01). Estimates of ASA classification, higher cancer stage, age, and lymph ratio were negative.

We also fitted a Cox proportional-hazards regression model, but we do not interpret the results here, since the proportional-hazards assumption was violated.

We used the selected covariates and modeled the log survival time with expectile regression. We used IPCR weights to account for right-censoring. Age, number of examined lymph nodes, and lymph ratio were modeled with P-splines. The asymmetry levels were chosen as in Section 3. Results from the expectile regression can be seen in Table 7, Figures 5 and 6. Table 7 shows the estimated regression coefficients. The estimates for the 0.5-expectile differed slightly from the Weibull model estimates and had larger standard errors. The estimate for chemotherapy was 1.1 (95% CI: 0.52 to 1.67) and the estimate for ASA score (reference: smaller than 3) was −0.45 (CI: −1.02 to 0.12). The association with a higher cancer stage was negative. The associations with chemotherapy and ASA classification were estimated differently for patients with shorter survival times. The estimate for ASA score decreased to as low as −1.12 (r = 0.05, CI: −1.97 to −0.27). The estimate for chemotherapy increased to 1.38 (CI: 0.61 to 2.16) at the 0.1-expectile level and up to 1.94 (CI: 1.06 to 2.82) at the 0.01-expectile level. Since prior research suggested a maximum at the 7.5%-expectile, the estimates for the
0.01- and 0.02-expectiles are larger than expected. It should be noted, however, that the standard errors for the lower part of the distribution were elevated. The results can be explained to doctors and patients by talking about worst, average, and best-case scenarios. In the worst-case scenario, if the patient has a rather short survival time, chemotherapy is associated with the largest benefits.

Figure 6 shows the predicted expectiles of the log survival time as a function of the number of examined lymph nodes. There were similar trends for most expectile level. An increase in the number of examined lymph nodes was positively associated with longer survival times up to 51 lymph nodes. The strongest positive association was at 40 lymph nodes. Examining more than 51 lymph nodes was negatively associated with survival time. However, only around 4% of the patients had 51 lymph node examinations or more.

In Figure 7 we show the estimated expected residual life (ERL) of patients with chemotherapy treatment against patients without. The ERL was calculated for the log event time and transformed back to the event time by using the exponential function. This transformation was done to simplify interpretation. The ERL demonstrated a positive association between survival time and chemotherapy. At 9 months survival, ERL was estimated to be 1.23 years with chemotherapy
and 0.48 years without. After 2 years, the expected residual life was close to 0 (around 28 days) for patients who did not receive chemotherapy. Patients which received chemotherapy still had an ERL of 0.67 after 2 years, the ERL was close to 0 after 3.3 years. After these minima, the ERL increased again (not shown due to scarcity of the data). Differences were largest at 0.6 and 1.6 years, although the lack of smoothness of the curves should be taken into account.

5 DISCUSSION

We proposed the use of IPC weights to extend expectile regression to right-censored data. If the true weights are known, asymptotics follow from M-quantile regression. For unknown IPC weights, we have used the Kaplan-Meier estimator and the plug-in principle to adapt point estimates as well as estimates of the covariance matrix.

Our extensive simulation study has shown that the IPC estimator often underestimates regression parameters. Estimates for the upper tail of the distribution often showed the largest absolute bias. We also analyzed coverage of corresponding confidence intervals. The desired confidence was not achieved in most cases, especially for the upper expectiles. In our opinion the poor coverage in scenarios 1 to 3 stemmed from the bias of the estimator, rather than the standard error estimation. The coverages were much closer to the desired level in scenario 4, where the bias was reduced drastically. The event time was bounded in this scenario. For our estimator, we assumed that event times had a nonzero chance to be uncensored. If the chance was close to zero for some observations from the upper part of the distribution (scenarios 1 to 3), we observed a negative bias. The bias was largest for upper expectiles which place more weight on large observations.

We analyzed survival times of colon cancer patients with expectile regression. We found that there was a positive association between chemotherapy and overall survival for the average patient. For the lower expectiles the positive association was stronger. We analyzed the effect of an increased number of examined lymph nodes. In the case of lymphogenic metastasis of the colon carcinoma, the AWMF S3 guideline for colorectal carcinoma describes that metastatic spread happens according to a regular metastasis pattern, namely initially longitudinally, on both sides of the tumor, into the paracolic lymph nodes, then further down to the intermediate lymph nodes and along the radial arteries up to the central lymph nodes at the trunk of the supplying arteries. Therefore, the AWMF S3 guideline recommends resection of at least 12 lymph nodes. It also states that the more lymph nodes are harvested and histopathologically examined, the better. Our results showed the strongest positive association at 40 lymph nodes.

In our analyses, we also illustrated the advantages of expectile regression. In contrast to an AFT model, we did not assume a parametric distribution for the response which could potentially be a bad fit. Instead we estimated a rich set of expectiles to construct a conditional empirical distribution function for the response. We constructed semiparametric predictors in combination with quadratic penalties for strong flexibility in the inclusion and estimated effects of the covariates. Quantile regression as the main nonparametric alternative to expectile regression offers much more limited options to model covariates. Current extensions of quantile regression for right-censored data only allow for a simple linear model structure. Expectile regression also offers an easy way to calculate a tail expectation, or in this

![Figure 7](https://wileyonlinelibrary.com)
scenario the expected residual lifetime. We would like to recommend this as an intuitive alternative to the hazard rate.

The use of noncrossing expectile regression estimates would be an obvious way to further develop our method. We have seen some crossing expectiles in our main analysis as shown in Figure 5. These might be caused mainly by the difficult estimation of smoothing parameters in the tails. Still, in an additive model some crossings in covariate effects might be intentional while the overall predictors are still monotonously increasing. Nevertheless, the introduction of IPC weights to the expectile bundle model⁶ would be intriguing. Although it is not straightforward how the IPC would be included in the estimation of the scale part of the location-scale model. We think that further research in this direction might be warranted. Extending expectile regression with IPC was our initial approach in our project on distributional regression for censored data. In the next steps, the IPC method would benefit from further research. Future research could focus on reducing the bias of the point estimator. A correction is particularly necessary for situations where the probability for events is low at the upper tail of the distribution. IPCRR weights lead to reduced bias in comparison with IPCKM, since otherwise lost weights were added to the last observation. One approach might include a correction time for the last observation when it is censored. The IPC idea itself is so universal that it may also be implemented into other methods beyond mean regression.

Further, we will look into likelihood-based approaches that will also allow us to improve the quantification of censored observations for the estimates. We hope that there will also be a possibility within this to increase the flexibility of a quantile regression method for censored data.

ACKNOWLEDGEMENTS
The authors thank the involved pathologists. Financial support from the German Research Foundation (DFG) grant SO1313/1-1 project “Distributional Regression for Time-to-Event Data” is gratefully acknowledged. The authors also appreciate the helpful comments of two anonymous reviewers which led us to substantially improve the article.

DATA AVAILABILITY STATEMENT
The data that support the findings of this study are available on request from the corresponding author. The data are not publicly available due to privacy or ethical restrictions.

ORCID
Alexander Seipp  https://orcid.org/0000-0002-1496-4818
Verena Uslar  https://orcid.org/0000-0003-3252-2076

REFERENCES
1. Rigby RA, Stasinopoulos DM. Generalized additive models for location, scale and shape. *J Royal Stat Soc Ser C (Appl Stat)*. 2005;54(3):507-554. https://doi.org/10.1111/j.1467-9876.2005.00510.x.
2. Koenker R, Bassett G. Regression quantiles. *Econometrica*. 1978;46(1):33-50.
3. Newey WK, Powell JL. Asymmetric least squares estimation and testing. *Econometrica*. 1987;55(4):819-847.
4. Sobotka F, Kneib T. Geodlicative expectile regression. *Comput Stat Data Anal*. 2012;56(4):755-767. https://doi.org/10.1016/j.csda.2010.11.015.
5. Hall WJ, Wellner JA. *Estimation of Mean Residual Life*. New York, NY: Springer International Publishing; 2020:169-189.
6. Schnabel SK, Eilers PHC. A location-scale model for non-crossing expectile curves. *Stat*. 2013;2(1):171-183. https://doi.org/10.1002/sta4.27.
7. Schulze-Waltrup L, Sobotka F, Kneib T, Kauermann G. Expectile and quantile regression – David and Goliath? *Stat Model*. 2015;15(5):433-456.
8. Portnoy S. Censored regression quantiles. *J Am Stat Assoc*. 2003;98(464):1001-1012. https://doi.org/10.1198/016214503000000954.
9. Peng L, Huang Y. Survival analysis with quantile regression models. *J Am Stat Assoc*. 2008;103(482):637-649. https://doi.org/10.1198/016214508000000355.
10. Tsiatis A. *Semiparametric Theory and Missing Data*. New York, NY: Springer-Verlag; 2006.
11. Koul H, Susarla V, Ryzin JV. Regression analysis with randomly right-censored data. *Ann Stat*. 1981;9(6):1276-1288.
12. Zhou M. M-estimation in censored linear models. *Biometrika*. 1992;79(4):837-841.
13. Huang J, Ma S, Xie H. Least absolute deviations estimation for the accelerated failure time model. *Stat Sin*. 2007;17(4):1533-1548.
14. German Cancer Society, German Cancer Aid, & AWMF (German Guideline Program in Oncology) S3-Guideline colorectal cancer, long version 2.1; 2019. http://www.leitlinienprogrammonkologie.de/leitlinien/kolorektales-karzinom/. Accessed February 24, 2021.
15. Brandi G, De Lorenzo S, Nannini M, et al. Adjuvant chemotherapy for resected colorectal cancer metastases: literature review and meta-analysis. *World J Gastroenterol*. 2016;22(2):519.
Seipp A, Uslar V, Weyhe D, Timmer A, Otto-Sobotka F. Weighted expectile regression for right-censored data. *Statistics in Medicine*. 2021;40(25):5501-5520. https://doi.org/10.1002/sim.9137

**APPENDIX A. UNIVARIATE EXPECTILES IN TIME-TO-EVENT ANALYSIS**

Univariate expectiles $\mu(\tau)$ of a distribution $T$ are defined by the equation

$$\tau = \frac{K(\mu(\tau)) - \mu(\tau)F(\mu(\tau))}{2[K(\mu(\tau)) - \mu(\tau)F(\mu(\tau))] + \mu(\tau) - \mu(0.5)},$$

where $K(t)$ is the (first) partial moment of $T$ and $F(t)$ is the cdf. The partial moment is defined by $K(t) = \int_{-\infty}^{t} x \, dF(x)$. Partial moments for several distributions were given in Winkler et al.\(^{30}\) including the gamma distribution. We present partial moments of two distributions that are frequently used in survival analysis.
(1) Let $T$ follow a Weibull distribution with density $f(t) = \lambda k (\lambda t)^{k-1} \exp\left(- (\lambda t)^k\right), \lambda \in \mathbb{R}^+, k \in \mathbb{R}^+$. The partial moment for $t > 0$ is

$$K(t) = \frac{1}{\lambda} \gamma\left(1 + \frac{1}{k}(\lambda t)^k\right),$$

where $\gamma(a, b) = \int_0^b x^{a-1} \exp(-x) \, dx$.

(2) Let $T$ follow a logistic distribution with $f(t) = \exp\left(\frac{\alpha - t \beta}{1 + \exp\left(\frac{\alpha - t \beta}{\beta}\right)}\right), \alpha \in \mathbb{R}, \beta \in \mathbb{R}^+$. The partial moment is

$$K(t) = \frac{t \exp\left(\frac{t}{\beta}\right)}{\exp\left(\frac{\alpha}{\beta}\right) + \exp\left(\frac{t}{\beta}\right)} - \beta \log\left(\exp\left(\frac{\alpha}{\beta}\right) + \exp\left(\frac{t}{\beta}\right)\right).$$

We can use the partial moments to calculate expectiles if the theoretical distribution is known. In Section 3 we have carried out a simulation study and specified different event time distributions. The theoretical expectiles served as a benchmark for predicted expectiles given by expectile regression.

**APPENDIX B. ADDITIONAL TABLES**

Tables B1 and B2 present further results for point estimates in base scenarios of the simulation study.

| \(\tau\) | Bias | MSE |
|----------|------|-----|
|          | Full data | IPCKM | IPCRR | Full data | IPCKM | IPCRR |
| 0.01     | 0.04   | -0.05 | 0.01  | 0.13      | 0.16  | 0.16  |
| 0.02     | 0.02   | -0.08 | -0.01 | 0.09      | 0.13  | 0.12  |
| 0.05     | 0.01   | -0.11 | -0.03 | 0.06      | 0.10  | 0.09  |
| 0.1      | 0.01   | -0.13 | -0.04 | 0.04      | 0.10  | 0.08  |
| 0.2      | 0.00   | -0.16 | -0.06 | 0.03      | 0.11  | 0.08  |
| 0.5      | -0.00  | -0.24 | -0.11 | 0.03      | 0.17  | 0.11  |
| 0.8      | -0.01  | -0.38 | -0.21 | 0.03      | 0.31  | 0.20  |
| 0.9      | -0.01  | -0.50 | -0.31 | 0.04      | 0.46  | 0.29  |
| 0.95     | -0.02  | -0.64 | -0.45 | 0.06      | 0.67  | 0.43  |
| 0.98     | -0.03  | -0.87 | -0.69 | 0.09      | 1.07  | 0.75  |
| 0.99     | -0.05  | -1.08 | -0.90 | 0.13      | 1.51  | 1.12  |
TABLE B2  Bias and MSE in base scenario 4 for \( n = 200 \)

| \( \tau \) | Bias | MSE |
| --- | --- | --- |
| | Full data | IPCKM | IPCRR | Full data | IPCKM | IPCRR |
| 0.01 | 0.01 | 0.01 | 0.01 | 0.00 | 0.01 | 0.01 |
| 0.02 | 0.01 | 0.00 | 0.01 | 0.00 | 0.01 | 0.01 |
| 0.05 | 0.01 | 0.00 | 0.01 | 0.01 | 0.01 | 0.01 |
| 0.1 | 0.01 | -0.00 | 0.00 | 0.01 | 0.01 | 0.01 |
| 0.2 | 0.00 | -0.01 | -0.00 | 0.01 | 0.02 | 0.02 |
| 0.5 | -0.00 | -0.03 | -0.01 | 0.01 | 0.02 | 0.02 |
| 0.8 | -0.01 | -0.04 | -0.03 | 0.01 | 0.03 | 0.03 |
| 0.9 | -0.01 | -0.05 | -0.03 | 0.01 | 0.03 | 0.03 |
| 0.95 | -0.01 | -0.06 | -0.04 | 0.01 | 0.03 | 0.03 |
| 0.98 | -0.01 | -0.07 | -0.06 | 0.00 | 0.04 | 0.03 |
| 0.99 | -0.01 | -0.08 | -0.07 | 0.00 | 0.04 | 0.03 |