

Reliability with multiple causes of failures: Modeling and practice through a case study on ultrasound probes for medical imaging

Rossella Berni^{1,*} , Francesco Bertocci² , Alessandro Magrini³ , and Nedka D. Nikiforova³ 

¹ Department of Statistics Computer Science Applications “G.Parenti”, University of Florence, Viale Morgagni, 59, 50134, Florence, Italy

² R&D Global Transducer Technology, Esaote Spa, Via di Caciolle, 15, 50127 Florence, Italy

³ Department of Statistics Computer Science Applications “G. Parenti”, University of Florence, Italy

Received: 30 January 2024 / Accepted: 14 July 2024

Abstract. In this paper, we deal with statistical modeling and a related case study for reliability when multiple failure causes are present. At first, we present in detail two main approaches for competing risk modeling, e.g. the Cox Proportional Hazards model, and the Fine & Gray model. In both models, we consider the inclusion of random effects, a non-trivial issue in this context, especially from the practical point of view. Following, we deal with advanced statistical models to compare the causes of failure, providing extremely useful information for production managers. To perform a useful study for practitioners, statistical modeling is illustrated through an empirical example related to ultrasound probes for medical imaging. The main theory is briefly presented comprehensively, while particular emphasis is given to data structure for model estimation and interpretation of the results, highlighting methodological comparisons and practical differences. Details related to two statistical software are also provided. Furthermore, reliability modeling could be efficiently applied by practitioners and engineers to solve similar technical problems.

Keywords: Competing risks / random effects / Weibull models / ultrasound transducers

1 Introduction

Nowadays, statistics is crucial for the technological and engineering fields. Through implementing methods and modeling, statistical responses are fundamental for supporting management throughout the decisional steps. In this way, using statistics in these fields requires: i) a good knowledge of the engineering issues and aspects and ii) an adequate knowledge of statistical methods. These latter two key points imply a deep and open interdisciplinary perspective from both points of view: statisticians and engineers.

In this paper, we focus on the application of specific statistical models for reliability, and we attempt to give some useful guidelines for practitioners. To this end, reliability modeling is illustrated through a real case study based on a simulated data-set starting from real data. Furthermore, it aims to provide a practical guide for analyzing similar technological problems in the reliability field. The contribution is further discussed with the

engineers involved in the real production process, and the results are illustrated in detail by considering the statistical meanings (and interpretations) of the estimated models and coefficients as well as the technical implications. Moreover, the specific parametrization for each applied model is carefully illustrated.

When considering failures and the type (cause) of failures, the statistical models can be divided into two distinct parts: i) the competing risks models, where each type of failure is modeled separately; ii) the Weibull models, where the causes of failure are treated through a covariate in order to evaluate the risk of one cause of failure with respect to another one (within the same type of ultrasound probe).

Furthermore, two dedicated software are applied and compared when possible: R for Statistical Computing [1], package `coxme` [2], and SAS (Statistical Analysis System, Windows Platform, Version 9.4), procedures `LIFEREG` and `PHREG`.

The real case-study is focused on the reliability of ultrasound probes (USPs) for medical imaging. More specifically, two different types of USPs are studied: probe-a and probe-b. The USPs may fail due to four causes of

* Corresponding author: rossella.berni@unifi.it

failure, described in detail in [Section 2.2](#). The two modeling approaches are as follows. First, we consider the two main competing risks models, that is, the Cox Proportional Hazards (PH) model [\[3\]](#) and the Fine & Gray model [\[4\]](#), with the inclusion of random effects. Secondly, Weibull and Cox's statistical models are estimated for each USP type, where the cause of failure, at four levels, is included as a covariate.

It must be noted that in both situations, e.g. competing risks and comparison between cause of failures, the Cox PH model is applied. Nevertheless, in [Section 3](#) the standard (classical) Cox PH model is adjusted to treat competing risks when the cause-specific hazard function is modeled.

Satisfactory results are obtained by applying both approaches by also considering the different information they provide in relation to specific technical aims.

In what follows the organization of the manuscript, with a short summary of contents:

- In [Subsection 2.1](#) we report a brief literature review and technical details on USPs; in [Subsection 2.2](#) we describe in detail the data we are dealing with the two USP types, including the specific simulated random effects, and the four causes of failure.
- In [Section 3](#), we report the main theory and the model results. More specifically, [Subsection 3.1](#) includes the definition and a detailed description of the main reliability functions. Moreover, the key concept of censoring is explained and illustrated through examples. In addition, the two main competing risk models are described. The main reliability functions (cumulative and probability density functions, survival and hazard functions) are schematically reported.
- In [Subsection 3.3](#) the competing risks model results are illustrated for the USPs in detail. An illustration of the dataset arrangement to apply these models is also included.
- In [Subsection 4](#), the Weibull and the Cox PH models to evaluate the risk of one cause of failure to another one are illustrated. Model results and HRs are reported; moreover, HRs are shown in detail by specifically considering different distributions for the survival time, e.g. Weibull and Exponential ones.

Discussion and final remarks end the manuscript.

2 The ultrasound probes: technical details and data description

In [Section 2.1](#), a brief description of the USPs is given through a panoramic view into the recent specific literature; moreover, [Section 2.2](#) contains a detailed description of real as well as simulated data.

2.1 Technical details on USPs and related literature

Ultrasound diagnostic technology has increasingly become important in imaging and image-guided interventional procedures by using the USPs. These latter transmit waves and receive ultrasonic echoes reflected from the patient tissue, and they convert the received pulses into electrical

signals, that are processed to generate a diagnostic image [\[5\]](#). In [Figure 1](#) a general scheme of a USP considering the main parts is shown.

The general structure of the USPs is characterized by piezoelectric array elements that are the source and the receiver for ultrasound waves. The structure also includes the backing, that provides the mechanical support for the piezoelectric material as well as the mitigation of acoustic noise effects. The acoustic matching layer improves the transfer of energy to the lens. This latter focuses on the ultrasound beam, and it is in direct contact with the patient [\[6\]](#). The electrical interconnections play a crucial role by considering the direct connection between the probe and the ultrasound system, through the interconnections board and the signal cable. The plastic-based housings that contain the ultrasound transducer separates the internal electronics from the outside: their shape and the sealing of each other are fundamental for the safety of the patient and for the probe's reliability over time.

Ad hoc electroacoustic tests provide measurements and qualitative tests to recognize if the USPs are damaged and worn out. Failures are categorized as minor or major faults that are mainly classified as lens damage, image reverberation, and low uniformity, physical damage of the USPs [\[7\]](#). Furthermore, the influence of bonding delaminations in ultrasound transducer can lead to the deterioration of the electroacoustic performance due to the loss of adhesion among some constitutive elements, such as the active element, the backing, or the matching layers [\[8\]](#).

It is relevant to note that, in general, all transducers must be cleaned and disinfected after each use. Cleaning is an important procedure that must be performed before disinfecting the transducer. Using an inappropriate disinfectant can damage the transducer. In fact, the manual for each probe lists compatible disinfectants; the use of a disinfectant not present within the list can cause damage to the USP structure.

Probe life may last longer if you remove the gel and cover it immediately after probe use since it prevents lens discoloration. Moreover, tissues/kerchiefs that prevent lens wear are usually recommended to remove the gel. Using a probe in an environment outside of specifications can lead to discoloration of some components, such as the cable; nevertheless, the color change does not affect the probe's performance or use.

The majority of probe faults can be detected by visual inspection and in-air reverberation assessment, due to the array damages, cable faults and loss of the electrical capacitance of one or several elements [\[9\]](#). The latter negatively affects the accuracy of the image quality and the Doppler-derived diagnostic information [\[10\]](#). In-air reverberation images-based methodology is used in [\[11\]](#) to detect failures in linear and curvilinear ultrasound transducers. In [\[12\]](#) a method for automatic detection of faulty linear array ultrasound transducers, based on the uniformity assessment of clinical images is discussed. In [\[13\]](#) the study, based on the Failure Mode Effect and Criticality Analysis (FMECA) classifies the impact of the probe failures on medical diagnoses from no impact to inability to carry out diagnostic work, by considering the maintenance efficiency, the costs, and the related risks.

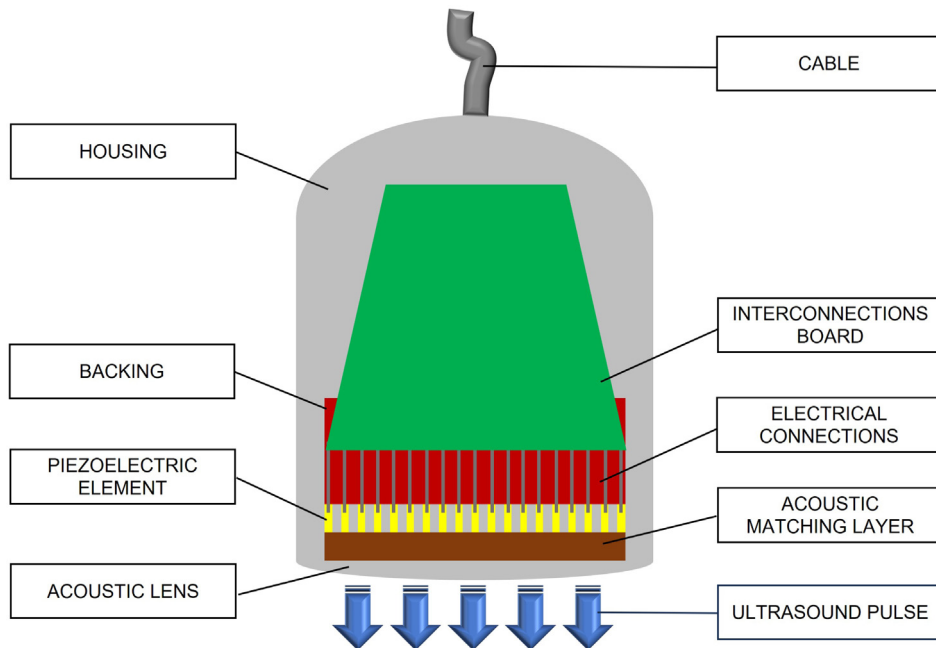


Fig. 1. A scheme of a generic USP: an ultrasound transducer inside the housing.

In [14] the acceptance testing is aimed to confirm that ultrasound equipment works according to the manufacturer's specifications and it also provides a reference baseline to test that the equipment performance remains constant over time.

2.2 Data description

The application of statistical modeling for reliability is illustrated through a case study based on real data on USPs for medical imaging produced by an international Company in the biomedical sector. A data-set is then simulated on the basis of real data on USPs by reproducing the main characteristics of starting data, and also including fixed and random variables.

In what follows the simulated data-set, reproducing the main characteristics of the real data, is described. The real USPs are sold between 1st January 2021 and 31st March 2022; the main differentiation is the USP type (G): probe-a (g_a) or probe-b (g_b). Probe-a has a flat array and appearance, with a specific application in musculoskeletal medical diagnosis. Probe-b has a curved linear array that allows for a wide field of view and is used for abdominal and obstetrics applications. Moreover, the internal parts of the linear US probe have lower thickness relating to a convex USP. More specifically, the layers of the multi-layer structure of the linear USPs are thinner, and therefore, the failure occurrence can be higher. For example, the lens in a linear USP is thin, and consequently, it can be more subject to failure with respect to the lens of the convex one. Furthermore, the different uses of the linear USP, e.g. for musculoskeletal medical diagnosis, imply that this USP is influenced by major pressure and a greater rubbing distance when performing the echo-imaging test (<https://www.esaote.com/>).

The duration in days (T) is available for each probe. Furthermore, and considering the USP lifetime in use, the number of ultrasonic probes that are the subject of notifications, e.g. subject to a failure, is analyzed and verified monthly. The temporary window, considering the standard warranty, is 2 yr (the period/date starting from when the probe is sold to when the firm receives the notification). In contrast, it is not considered the duty cycle or the number of examinations. The worst case at the operation level is evaluated, more specifically, about 15 daily sessions six days a week. Each session is supposed to be about 15 min, but it can vary from patient to patient. Defining a more precise operating perimeter is complex because there are several factors, such as physician handling, use in Doppler mode (higher severity for a probe because it gets hot) or not, etc. Probe life can last longer if the recommended disinfection manual is followed to prevent lens discoloration and wear and tear.

Moreover, if a probe does not experience any failure within the period of observation, the duration is censored at 455 days. Otherwise, the cause of failure (C) is recorded as one among the four following causes:

- array or contact (c_1),
- imaging, lens or dark zone (c_2),
- cable, system, safety or interference (c_3),
- plastics, no information or unknown (c_4).

Moreover, we considered two random variables, defined as follows:

- the degree of process improvement (W) random variable, describing the occurrence of interventions towards a continuous improvement that is interventions aiming at optimizing the product in reliability, following the well-known idea of ongoing efforts for achieving high level of

Table 1. Description of variables contained in the simulated data-set.

Variable	Description
G	USP type: g_a (probe-a), g_b (probe-b).
T	Time to failure: number of days.
C	Cause of failure: c_0 (no failure), c_1 (array or contact), c_2 (imaging, lens or dark zone), c_3 (cable, system, safety or interference), c_4 (plastics, no information or unknown).
W	Degree of process improvement: w_0 (absent), w_1 (present).
Z	Degree of process deterioration: z_0 (absent), z_1 (present).

quality and reliability product performances. The random variable W assumes two possible values: absent (w_0) and present (w_1);

- the degree of process deterioration (Z) random variable, describing the occurrence of accidental conditions like noise and technical issues that may increase the failure rate; it assumes two possible values: absent (z_0) and present (z_1).

The random variables Z and W are conceived to represent several degrees of improvement and deterioration for the current state of the process. They are simulated with just two values (absence and presence), leading to four possible situations: both improvement and deterioration are absent ($W=w_0$ and $Z=z_0$), improvement is present without deterioration ($W=w_1$ and $Z=z_0$), deterioration is present without improvement ($W=w_0$ and $Z=z_1$), both improvement and deterioration are present ($W=w_1$ and $Z=z_1$). The simulated data-set was built based on the following procedure:

- 8308 probes are simulated; this amount of probes is equal to the total number of sold probes, for both USP types, in the observed period;
- for each simulated probe, the random variables W and Z are randomly assigned following a uniform distribution across the four situations (w_0, z_0), (w_0, z_1), (w_1, z_0), and (w_1, z_1);
- for each probe, the duration (in days) is simulated based on pre-specified failure rates for each quarter of the considered period. These failure rates depend on USP type, cause of failure, absence/presence of the process improvement (W) and absence/presence of the process deterioration (Z). In particular, failure rates were set equal to actual ones in the absence of both improvement and deterioration ($W=0$ and $Z=0$), lower than actual ones in the presence of improvement ($W=1$) and higher than actual ones in the presence of deterioration ($Z=1$);
- for each probe showing a duration greater than or equal to 455 days, the duration is censored at 455 days.

Table 1 provides the data-set description. In addition, a screenshot of ten randomly chosen records from the data-set is displayed in Figure 2 to illustrate the data structure and input accordingly.

3 Main theory and results

This Section is organized as follows: in Subsection 3.1 and 3.2 the basic theory on censoring, competing risk concepts and random effects are reported [3,4,15], while the corresponding model results are shown in Subsection 3.3.

3.1 Censoring types and basic functions in reliability

Let us start by defining some basic functions when dealing with reliability data. Let T be a continuous random variable describing the time to failure for a statistical unit. The probability distribution of T can be characterized by the following four main functions: the cumulative distribution function $F(t)$, the probability density function $f(t)$, the survival function $S(t)$ (also called the reliability function), and the hazard function $h(t)$. These main functions are defined as follows.

- **Cumulative distribution function** $F(t)$: the probability to fail before time t :

$$F(t) = P(T \leq t).$$

- **Probability density function** $f(t)$: the rate of failure at time t :

$$f(t) = \frac{d}{dt} F(t).$$

- **Survival function** $S(t)$: the probability to survive up to time t :

$$S(t) = P(T > t) = 1 - F(t).$$

- **Hazard function** $h(t)$: the rate of failure at a time t , given survival up to time t :

$$\begin{aligned} h(t) &= \lim_{\Delta t \rightarrow 0} \frac{P(t < T \leq t + \Delta t | T > t)}{\Delta t} = \frac{f(t)}{1 - F(t)} \\ &= \frac{f(t)}{S(t)}. \end{aligned}$$

A key-point in the reliability context is related to the type of censoring. The first type is the right censoring (Type I, Type II, and random), which is the most common one. It occurs when a statistical unit leaves the study before an event of interest occurs, or when the study ends before the occurrence of the event. Let T_i and C_i^* be the survival time and the censoring time for unit i ($i=1, \dots, n$) respectively. For each i -th unit, we observe (t_i, δ_i^*) where $t_i = \min(T_i, C_i^*)$, while δ_i^* is defined as follows:

$$\delta_i^* = \begin{cases} 1 & \text{if } T_i \leq C_i^* \\ 0 & \text{if } T_i > C_i^* \end{cases}.$$

ID	G	T	C	Z	W
577	1	455	0	1	0
4553	1	397	2	1	1
1015	2	455	0	0	0
1230	2	455	0	1	0
7897	1	84	4	0	1
1494	2	398	3	0	0
3353	1	455	0	1	0
3080	1	169	2	1	0
5870	1	184	2	0	1
270	1	455	0	0	1

Showing 1 to 10 of 8,308 entries, 6 total columns

Fig. 2. Data structure and input: ten randomly selected records from the simulated data-set.

More specifically, if $\delta_i^* = 0$, then the time to failure is censored, instead if $\delta_i^* = 1$ then the time to failure is completely known, e.g. no censoring occurs. Obviously, Type I right censoring occurs when the study ends at a predetermined period. At the end of this fixed period, statistical units which do not experience the event of interest are censored. Therefore, the number of censored cases is random, while the censoring time is fixed. In Type II right censoring, the study ends when a fixed number of events of interest occur. Therefore, with this type of censoring, the censoring times are random, while the number of censored cases is fixed. In the USPs case study, we are dealing with Type I right censored data. The USPs refer to the period January 1st 2021- March 31st 2022; therefore, those that do not experience any failure are censored. For further details on right censoring and left and interval types of censoring, refer to [15,16].

3.2 Outlined theory on competing risk models with random effects

Competing risks take place in studies in which a statistical unit (product/component) may experience several failure events (not a single failure). The main assumption in

competing risks setting is that the causes of failure are mutually exclusive, e.g. the failure by one cause precludes the occurrence of any other failure. Therefore, we could say that in the competing risk case, a statistical unit may experience only one cause of failure among a set of potential ones. Two main modeling approaches exist to deal with competing risks: the cause-specific hazard regression model, and the subdistribution hazard regression model. In the former, the Cox PH model [3] is opportunely adapted to study the cause-specific hazard function; that is, for a given event of interest, the Cox PH model is estimated, treating the remaining competing events as censored observations. The latter, developed by [4], takes the subdistribution hazard function into account. It denotes the instantaneous rate of failure due to a specific cause for those statistical units that do not experience failure or drop the study for other motivations.

The cause-specific hazard function $h_k(t)$ represents the instantaneous rate of failure due to the k -th cause conditioned to the survival until time t , and it is defined as follows:

$$h_k(t) = \lim_{\Delta t \rightarrow 0} \frac{P(t < T < t + \Delta t, C = c_k | T > t)}{\Delta t}. \quad (1)$$

Therefore, the standard Cox PH model for competing-risks data is defined as follows:

$$\mathbf{h}_k(t) = h_{0k}(t) \cdot \exp(\mathbf{X}\boldsymbol{\beta}) \quad k = 1, 2, 3, 4 \quad (2)$$

where $h_{0k}(t)$ is the unspecified baseline hazard function for cause c_k , \mathbf{X} is the model matrix of covariates, and $\boldsymbol{\beta}$ is the corresponding vector of unknown coefficients. Therefore, a separate Cox PH model should be estimated for each k -th cause. A fundamental assumption of model (2) is the proportional hazard, e.g. the hazard ratio of any covariate is hypothesized constant over time [16]. To this end, it is crucial to always check this assumption by using graphical and numerical methods [15,16]. Lastly, it must also be noted that in the Cox model, formula (2), the causes of failure are assumed to be independent. In the case study, this assumption implies, for example, a probe that failed due to an array fault (c_1) has the same chance of failing due to an imaging fault (c_2) as any other probe that experiences no failure.

The second modeling approach for competing risks data is related to the sub-distribution hazard regression model, also known as the Fine & Gray model [4]. In this case, the sub-distribution hazard function $\lambda_k(t)$ represents the instantaneous rate of failure due to the k -th cause conditioned to either (i) survival until time t , or (ii) failure before time t due to other causes, and it is defined as follows:

$$\lambda_k(t) = \lim_{\Delta t \rightarrow 0} \frac{P[t < T < t + \Delta t, C = c_k | T > t \cup (T < t \cap C \neq c_k)]}{\Delta t} \quad (3)$$

The Fine & Gray model with competing risks is expressed as the following:

$$\lambda_k(t) = \lambda_{0k}(t) \cdot \exp(\mathbf{X}\boldsymbol{\beta}) \quad k = 1, 2, 3, 4. \quad (4)$$

where $\lambda_{0k}(t)$ is the unspecified baseline sub-distribution hazard function for cause c_k , \mathbf{X} is the model matrix of covariates, and $\boldsymbol{\beta}$ is the corresponding vector of unknown coefficients. Therefore, differently from model (2), model (4) imposes the proportional hazard assumption on the sub-distribution hazard function. Moreover, it must be noted that the assumption of independent causes of failure is relaxed by model (4), [4]. In fact, they are equal to the log sub-distribution hazard function, instead of the log hazard function, for each cause of failure.

An important key point is including random effects in this field, also named frailties [17,18]. In the beginning, the term frailty has been introduced since the 70s to account for unobserved heterogeneity of survival data, [19]. The frailty term is generally modeled through an unobserved random effect that multiplicatively affects the hazard function. Such random effects, in general, are used to account for unobserved individual heterogeneity, correlation in clustered data, or for recurrent events, and to also explain issues related to lack of fit for example violations of the proportional hazard assumption [17,18,20]. In general, we may assume different distributions for the frailty term. More specifically, we may consider: i) the Laplace transform, ii) the infinitely divisible distributions, as the

Gamma, the inverse Gaussian, and the compound Poisson distributions, and iii) the Log-Normal distribution. In general, the choice of the frailty distribution is made by convenience. Balan and Putter [20] discuss in detail all the issues related to the choice of the frailty distribution, also reporting some useful practical considerations, primarily related to the Gamma and the Log-Normal distributions. More specifically, the Gamma distribution is the simplest and most widely used in this field; moreover, regardless of the frailty distribution at the baseline, it is also the limiting distribution of the frailty of long-time survivors [21]. However, as also discussed by Ripatti and Palmgren [22], the Gamma distribution may run into computational difficulties, especially when more complicated dependence frailty structures are assumed. To this end, the Log-Normal distribution is the most commonly assumed one, also considering that it involves an additive random effect on the same scale of the covariates [20,22,23].

Both models (2) and (4) can be extended to account for random effects. More specifically, the competing risks Cox PH model, including random effects, is defined as follows:

$$\mathbf{h}_k(t) = h_{0k}(t) \cdot \exp(\mathbf{X}\boldsymbol{\beta} + \mathbf{V}\boldsymbol{\zeta}) \quad k = 1, 2, 3, 4. \quad (5)$$

where $h_{0k}(t)$ is the unspecified baseline hazard function for cause c_k , \mathbf{X} and \mathbf{V} are the model matrices of fixed and random effects respectively, while $\boldsymbol{\beta}$ and $\boldsymbol{\zeta}$ are the corresponding vectors of fixed and random-effect coefficients. Similarly, the competing risks Fine & Gray model including random effects is expressed as follows:

$$\lambda_k(t) = \lambda_{0k}(t) \cdot \exp(\mathbf{X}\boldsymbol{\beta} + \mathbf{V}\boldsymbol{\zeta}) \quad k = 1, 2, 3, 4. \quad (6)$$

where $\lambda_{0k}(t)$ is the unspecified baseline sub-distribution hazard function for the cause c_k , \mathbf{X} and \mathbf{V} are the model matrices of fixed and random effects respectively, while $\boldsymbol{\beta}$ and $\boldsymbol{\zeta}$ are the corresponding vectors of fixed and random-effect coefficients. In general in this field, the random effects are usually assumed distributed considering three distributions: Gamma, inverse Gaussian or Log-Normal [20]. In the case study, we assume the Log-Normal distribution which corresponds to Normally distributed random effects on the scale of the covariates. The Log-Normal distribution is a valid and widespread choice, as already discussed in this section. Moreover, it is relevant to note that, for identifiability reasons, a finite mean must be assumed for the frailty distribution [24]; that is, for the Log-Normal distribution, a mean equal to zero is the most widely used choice [22,23,25].

In the following subsection 3.3, matrix \mathbf{X} contains the fixed variables C and G , while matrix \mathbf{V} contains the random variables W and Z , Table 1.

3.3 Cox's and Fine and Gray's model results

We illustrate the application of the competing risks model (5) to the simulated data-set. The model assumes that the log hazard function for each cause of failure c_k , $k=1,2,3,4$, is proportional to a linear combination of the values of the covariates. Therefore, the categorical covariates (the USP type G at fixed levels, the degree of process improvement

W , and the degree of process deterioration Z) must preliminarily undergo dummy coding. More precisely, if the covariate is dichotomous, then the dummy coding originates a binary variable taking value one if the covariate takes the reference value and zero otherwise. For instance, let us consider the USP type (G), which takes the values g_a and g_b . If the value g_a is chosen as a reference level, the dummy coding originates a binary variable that maps the value g_b into one and the value g_a into zero. In general, if the covariate is polytomous, the dummy coding originates one binary variable for each possible value except the reference one (e.g. the number of dummies must be equal to the number of degrees of freedom of the variable). Let us suppose, for example, that a third USP type g_c exists. In this case, two binary variables are created: the first one maps the value g_b into one, and the other two values (g_a and g_c) map into zero, while the second one maps the value g_c into one, and the other two values (g_a and g_b) map into zero.

In the data-set, all the categorical variables are dichotomous. Therefore a single dummy variable is required. For the USP type (G), we choose the probe-a (g_a) as a reference, therefore the dummy coding is represented by the variable \mathbf{d}_{g_b} that takes the value one if the USP type is g_b (probe-b), and zero otherwise. Instead, for the degree of process improvement (W), we choose the first degree (w_0 , absence) as a reference, therefore the dummy coding is represented by the variable \mathbf{d}_{w_1} , which assumes value one if the process improvement is present, and zero otherwise. Similarly, z_0 (absence) is chosen as the reference value for the degree of process deterioration (Z), leading to the dummy variable \mathbf{d}_{z_1} .

The competing risks Cox PH model applied to the simulated data-set, formula (5), is formulated assuming that the risk of failure is affected by:

- the USP type (g_a and g_b), with probe-a selected as a reference;
- the degree of process improvement (w_0 and w_1), with different strengths for each USP type (g_a and g_b);
- the degree of process deterioration (z_0 and z_1), with different strengths for each USP type (g_a and g_b).

Given the definition of the categorical variables, in the model formulation (5) the unknown $\boldsymbol{\beta}$ coefficients are fixed and associated with the USP type; the unknown vectors of random coefficients associated with process improvement and deterioration are $\boldsymbol{\gamma}$ and $\boldsymbol{\delta}$, with corresponding standard deviations denoted by σ and τ respectively. All the reliability models are estimated using the R for Statistical Computing [1], package `coxme` [2].

The competing risks Cox PH model including random effects, specifically for the USP case, is defined as shown in the following formula (7); the model results are reported in Table 2.

$$\mathbf{h}_k(t) = h_{0k}(t) \cdot \exp\left(\boldsymbol{\beta}_{k(b)} \mathbf{d}_{g_b} + \gamma_{k(a)} \mathbf{d}_{w_1} \mathbf{d}_{g_a} + \gamma_{k(b)} \mathbf{d}_{w_1} \mathbf{d}_{g_b} + \delta_{k(a)} \mathbf{d}_{z_1} \mathbf{d}_{g_a} + \delta_{k(b)} \mathbf{d}_{z_1} \mathbf{d}_{g_b}\right). \quad (7)$$

$$\begin{pmatrix} \gamma_{k(a)} \\ \gamma_{k(b)} \end{pmatrix} \sim N\left(\begin{pmatrix} 0 \\ 0 \end{pmatrix}, \begin{pmatrix} \sigma_k^2 & 0 \\ 0 & \sigma_k^2 \end{pmatrix}\right),$$

$$\begin{pmatrix} \delta_{k(a)} \\ \delta_{k(b)} \end{pmatrix} \sim N\left(\begin{pmatrix} 0 \\ 0 \end{pmatrix}, \begin{pmatrix} \tau_k^2 & 0 \\ 0 & \tau_k^2 \end{pmatrix}\right),$$

where:

- $h_k(t)$ is the hazard function for each cause c_k , $k=1,2,3,4$;
- $h_{0k}(t)$ is the (unspecified) baseline hazard function for cause c_k ;
- the model terms denoted by letters $\boldsymbol{\beta}$, $\boldsymbol{\gamma}$ and $\boldsymbol{\delta}$ are the vectors of unknown coefficients representing log HRs;
- $N(\cdot)$ indicates a Normal distribution parameterized through the expected value and the variance, where σ_k^2 is the variance for $\gamma_{k(a)}$ and $\gamma_{k(b)}$, while τ_k^2 is the variance for $\delta_{k(a)}$ and $\delta_{k(b)}$.

The model results (Tab. 2) allow us to make the following considerations:

- Probe-b shows lower failure rates due to all the causes (c_1, c_2, c_3, c_4) than probe-a, at a constant degree of process deterioration and process improvement. More specifically, the estimated HR of the cause c_1 is equal to 0.024, therefore the hazard of the cause c_1 for probe-b is $(0.024-1) \times 100 = -97.6\%$ more than the one for probe-a, or equivalently, the hazard of the cause c_1 for probe-a is $1/0.024 = 41.667$ times the one for probe-b. Analogously, the estimated HR of the cause c_2 is equal to 0.559, therefore the hazard of the cause c_2 for probe-a is $(0.559-1) \times 100 = -44.1\%$ than the one for probe-b, or equivalently, the hazard of the cause c_2 for probe-a is $1/0.559 = 1.789$ times the one for probe-b (+178.9%). When considering the cause c_3 , the estimated HR is equal to 0.378, then the hazard of the cause c_3 for probe-a is $(0.378-1) \times 100 = -62.2\%$ than the one for probe-b. The same result also applies for the cause c_4 , where the corresponding estimated HR indicates that the risk of the cause c_4 for probe-a is lower $((0.859-1) \times 100 = -14.1\%)$ than the one for probe-b. In addition, it must be noted that only the estimated coefficients related to c_1 and c_2 are statistically significant. For the c_3 cause of failure, the estimated coefficient, even though not significant, is relevant, while for the c_4 cause of failure, the corresponding p-value is not relevant but neither negligible.
- Process improvement shows a significantly negative impact on the hazard for the causes c_1, c_3 and c_4 relative to probe-a (coefficient $\gamma_{j(a)}$ for $j=1,3,4$), and a negative impact (with borderline significance) on the hazard for the cause c_3 relative to probe-b (coefficient $\gamma_{3(b)}$). More precisely, at a constant degree of process noise, the presence of process improvement decreases the risk of c_1, c_3 , and c_4 for probe-a by 58.9%, 84.5%, and 33.9%, respectively. On the other hand, the impact of process improvement appears higher for probe-b; in fact, the HR of the cause c_3 is equal to 0.042, corresponding to a decrease of 95.8%. This result suggests that probe-b is

Table 2. Model results by cause of failure and USP types – with random effects (formula (7)).

Cause of failure c_1	Estimate	SE	p -value	HR
$\beta_{1(a)}$	0.0000	–	–	–
$\beta_{1(b)}$	–3.7449	1.3078	0.0042	$\exp(-3.7449) = 0.024$
$\gamma_{1(a)}$	–0.8905	0.2520	0.0004	$\exp(-0.8905) = 0.411$
$\gamma_{1(b)}$	–0.3330	0.8330	0.6893	$\exp(-0.3330) = 0.719$
$\delta_{1(a)}$	1.4895	0.2996	0.0000	$\exp(1.4895) = 4.435$
$\delta_{1(b)}$	0.6290	1.1210	0.5747	$\exp(0.6290) = 1.876$
σ_1^2	0.8325	–	–	–
τ_1^2	1.7653	–	–	–
Cause of failure c_2	Estimate	SE	p -value	HR
$\beta_{2(a)}$	0.0000	–	–	–
$\beta_{2(b)}$	–0.5812	0.1977	0.0033	$\exp(-0.5812) = 0.559$
$\gamma_{2(a)}$	–0.1960	0.1162	0.0917	$\exp(-0.1960) = 0.822$
$\gamma_{2(b)}$	–0.0469	0.1433	0.7436	$\exp(-0.0469) = 0.954$
$\delta_{2(a)}$	0.3976	0.1350	0.0032	$\exp(0.3976) = 1.488$
$\delta_{2(b)}$	0.0222	0.1795	0.9014	$\exp(0.0222) = 1.022$
σ_2^2	0.0359	–	–	–
τ_2^2	0.0996	–	–	–
Cause of failure c_3	Estimate	SE	p -value	HR
$\beta_{3(a)}$	0.0000	–	–	–
$\beta_{3(b)}$	–0.9730	0.8306	0.2414	$\exp(-0.9730) = 0.378$
$\gamma_{3(a)}$	–1.8625	0.5933	0.0017	$\exp(-1.8625) = 0.155$
$\gamma_{3(b)}$	–3.1684	1.7152	0.0647	$\exp(-3.1684) = 0.042$
$\delta_{3(a)}$	1.3990	0.4829	0.0038	$\exp(1.3990) = 4.051$
$\delta_{3(b)}$	1.2654	0.7607	0.0962	$\exp(1.2654) = 3.545$
σ_3^2	11.9089	–	–	–
τ_3^2	2.0580	–	–	–
Cause of failure c_4	Estimate	SE	p -value	HR
$\beta_{4(a)}$	0.0000	–	–	–
$\beta_{4(b)}$	–0.1523	0.3437	0.6577	$\exp(-0.1523) = 0.859$
$\gamma_{4(a)}$	–0.4134	0.2056	0.0444	$\exp(-0.413) = 0.661$
$\gamma_{4(b)}$	–0.4003	0.2489	0.1077	$\exp(-0.4003) = 0.670$
$\delta_{4(a)}$	1.0576	0.2473	0.0000	$\exp(1.0576) = 2.880$
$\delta_{4(b)}$	0.6880	0.2861	0.0162	$\exp(0.6880) = 1.990$
σ_4^2	0.2195	–	–	–
τ_4^2	0.8522	–	–	–

more receptive to process improvement than probe-a. Moreover, it must be noted that the estimated coefficients for probe-a are statistically significant for all the causes of failure. For probe-b, the estimated coefficients are statistically significant at 10% level for the cause c_3 and almost statistically significant at 10% level for the cause c_4 . For causes c_1 and c_2 , the corresponding p-values are not significant, but neither are negligible.

- Process deterioration shows a relevant impact on the hazard for all the causes relative to probe-a (coefficient $\delta_{j(a)}$ for $j=1,2,3,4$), and a relevant impact on the hazard for the cause c_4 relative to probe-b (coefficient $\delta_{j(b)}$). In fact, at a constant degree of process improvement, the presence of process deterioration influences the hazard for all the causes of failure (c_1, c_2, c_3 , and c_4) related to probe-a; the hazard values are respectively 4.4, 1.5, 4.1, and 2.88 times the baseline situation. On the contrary, the impact of process noise appears lower for probe-b; in fact, the HR of cause c_4 is 1.990, which is 31% lower than probe-a's. This result indicates that the probe-b type is even more robust to processing noise than probe-a. When considering the corresponding p-values, it must be noted that they are statistically significant for probe-a for all the causes of failure. Lastly, for probe-b, the p-values are statistically significant at 10% level for the causes c_3 and c_4 , not significant for the cause c_1 , and negligible for the cause c_2 .

The competing risks Fine & Gray model including random effects, specifically defined for the USP case, is defined as follows:

$$\lambda_k(t) = \lambda_{0k}(t) \cdot \exp\left(\beta_{k(b)} \mathbf{d}_{g_b} + \gamma_{k(a)} \mathbf{d}_{w_1} \mathbf{d}_{g_a} + \gamma_{k(b)} \mathbf{d}_{w_1} \mathbf{d}_{g_b} + \delta_{k(a)} \mathbf{d}_{z_1} \mathbf{d}_{g_a} + \delta_{k(b)} \mathbf{d}_{z_1} \mathbf{d}_{g_b}\right), \quad (8)$$

$$\begin{pmatrix} \gamma_{k(a)} \\ \gamma_{k(b)} \end{pmatrix} \sim N\left(\begin{pmatrix} 0 \\ 0 \end{pmatrix}, \begin{pmatrix} \sigma_k^2 & 0 \\ 0 & \sigma_k^2 \end{pmatrix}\right)$$

$$\begin{pmatrix} \delta_{k(a)} \\ \delta_{k(b)} \end{pmatrix} \sim N\left(\begin{pmatrix} 0 \\ 0 \end{pmatrix}, \begin{pmatrix} \tau_k^2 & 0 \\ 0 & \tau_k^2 \end{pmatrix}\right),$$

where:

- $\lambda_{0k}(t)$ is the (unspecified) baseline subdistribution hazard function;
- the vectors of the model coefficients, β , γ and δ , have the same meaning as in the Cox PH model;
- HR interpretation is equal in both models, nevertheless in the Fine & Gray model the reference is the subdistribution hazard function $\lambda_k(t)$ rather than the hazard function $h_k(t)$. Finally, when considering the p-values related to the estimated coefficients, the same comments reported for the Cox PH model also apply for the Fine & Gray model.

The model results, formula (8), are shown in Table 3. We note a substantial similarity with the results obtained through the application of model (7). Therefore, we can conclude that the assumption of independent causes of failure, postulated by both models, is confirmed.

4 To compare causes of failure: advanced modeling

In this Section, we illustrate two further reliability models, e.g. the Weibull model and the Cox PH model [3,15,16], that allow for comparing the causes of failure and evaluating the risk of one cause of failure to another taken as a reference. At first, we report the basic theory for the Weibull model in Subsection 4.1. Following, we illustrate the model results by each USP type (Subt. 4.2). It must be noted that, differently from the reliability models reported in Section 3, we do not deal with competing risks through this modeling approach. Therefore, in this Section, we pursue a completely different aim than competing risk modeling. Furthermore, for both models i) a covariate is included to evaluate the cause of failure; ii) no random effects are included in this case.

4.1 The basic theory of the Weibull model

The Weibull model belongs to the parametric class of survival models, in which a particular shape for the hazard rate is specified according to the assumed distribution for time. The Weibull model is characterized by one location parameter α and one shape parameter γ . The shape of the hazard rate is assumed to be monotonic and determined by the shape parameter γ : decreasing when $\gamma < 1$, increasing when $\gamma > 1$, and flat when $\gamma = 1$. In this case, the Weibull model reduces to the Exponential model [16]. The cumulative distribution function $F_W(t)$ and the probability density function $f_W(t)$ for the Weibull model are defined as follows:

$$F_W(t) = 1 - \exp(-\alpha t)^\gamma,$$

$$f_W(t) = \alpha \gamma (\alpha t)^{\gamma-1} \exp(-\alpha t)^\gamma.$$

Moreover, the survival function $S_W(t)$ and the hazard function $h_W(t)$ are the following:

$$S_W(t) = \exp(-\alpha t)^\gamma,$$

$$h_W(t) = \gamma \alpha (\alpha t)^{\gamma-1}.$$

In a general expression, an Accelerated Failure Time (AFT) model considers the following relation between the dependent variable $\mathbf{Y} = \log \mathbf{T}$ and the covariates:

$$\log \mathbf{T} = \mathbf{X} \beta + \varepsilon. \quad (9)$$

Table 3. Model results by cause of failure and USP types – with random effects (formula (8)).

Cause of failure c_1	Estimate	SE	p -value	HR
$\beta_{1(a)}$	0.0000	–	–	–
$\beta_{1(b)}$	–3.7301	1.3009	0.0041	$\exp(-3.7301) = 0.024$
$\gamma_{1(a)}$	–0.8712	0.2516	0.0005	$\exp(-0.8712) = 0.418$
$\gamma_{1(b)}$	–0.3216	0.8198	0.6948	$\exp(-0.3216) = 0.725$
$\delta_{1(a)}$	1.4604	0.2992	0.0000	$\exp(1.4604) = 4.308$
$\delta_{1(b)}$	0.6131	1.1081	0.5801	$\exp(0.6131) = 1.846$
σ_1^2	0.8014	–	–	–
τ_1^2	1.7108	–	–	–
Cause of failure c_2	Estimate	SE	p -value	HR
$\beta_{2(a)}$	0.0000	–	–	–
$\beta_{2(b)}$	–0.5746	0.1950	0.0032	$\exp(-0.5746) = 0.563$
$\gamma_{2(a)}$	–0.1753	0.1126	0.1195	$\exp(-0.1753) = 0.839$
$\gamma_{2(b)}$	–0.0413	0.1367	0.7627	$\exp(-0.0413) = 0.960$
$\delta_{2(a)}$	0.3675	0.1332	0.0058	$\exp(0.3675) = 1.444$
$\delta_{2(b)}$	0.0172	0.1756	0.9219	$\exp(0.0172) = 1.017$
σ_2^2	0.0307	–	–	–
τ_2^2	0.0873	–	–	–
Cause of failure c_3	Estimate	SE	p -value	HR
$\beta_{3(a)}$	0.0000	–	–	–
$\beta_{3(b)}$	–0.9476	0.8244	0.2504	$\exp(-0.9476) = 0.388$
$\gamma_{3(a)}$	–1.8374	0.5932	0.0020	$\exp(-1.8374) = 0.159$
$\gamma_{3(b)}$	–3.1527	1.7064	0.0647	$\exp(-3.1527) = 0.043$
$\delta_{3(a)}$	1.3520	0.4812	0.0050	$\exp(1.3520) = 3.865$
$\delta_{3(b)}$	1.2428	0.7531	0.0989	$\exp(1.2428) = 3.465$
σ_3^2	11.7400	–	–	–
τ_3^2	1.9594	–	–	–
Cause of failure c_4	Estimate	SE	p -value	HR
$\beta_{4(a)}$	0.0000	–	–	–
$\beta_{4(b)}$	–0.1374	0.3432	0.6890	$\exp(-0.1374) = 0.872$
$\gamma_{4(a)}$	–0.3930	0.2043	0.0545	$\exp(-0.3930) = 0.675$
$\gamma_{4(b)}$	–0.3897	0.2464	0.1138	$\exp(-0.3897) = 0.677$
$\delta_{4(a)}$	1.0287	0.2469	0.0000	$\exp(1.0287) = 2.797$
$\delta_{4(b)}$	0.6833	0.2854	0.0167	$\exp(0.6833) = 1.980$
σ_4^2	0.2059	–	–	–
τ_4^2	0.8185	–	–	–

By considering formula (9), a different reliability model can be obtained according to the distribution we specify for the vector of residuals ε . For example, if ε is assumed to follow a Normal distribution (which is equivalent to assuming that time \mathbf{T} follows a Lognormal distribution), the Lognormal AFT model is obtained. Similarly, if ε follows a logistic distribution, the log-logistic AFT is obtained [16]. For this specific case-study, we decided to choose the Weibull model, as defined in the following AFT model specification:

$$\log \mathbf{T} = \mathbf{X}\boldsymbol{\beta} + \sigma\varepsilon. \quad (10)$$

In formula (10), the dependent variable is the logarithm of the survival time vector \mathbf{T} , \mathbf{X} is the matrix of values for the covariates, with 1s in the first column; moreover, the following two distributions are assumed: the Standard Extreme Value distribution for the error ε , and the Weibull distribution for the survival time \mathbf{T} . The Exponential distribution is another possible choice for the survival time, equating to the Weibull one when the shape parameter γ is equal to one.

4.2 To compare causes of failure: Weibull's and Cox's model results

In this Subsection, we illustrate the model results obtained for the USPs to compare the causes of failure. Therefore, we estimate two models, one for each USP type (probe-a and probe-b), including the covariate "cause of failure," at four levels (the categorical variable C , see Sect. 2.2).

It must be noted that the data-set should be arranged in an extended format. In Figure 3, we report an example of the extended data-set for the two USPs (counterfeit data for privacy). We can observe (Fig. 3) that we have four column vectors. The first column vector (**id**) is related to the identification of each USP. We have four rows for each USP since we have four causes of failure. The second column vector is related to the survival time (**time** in days). The third column vector is related to the variable "cause of failure." Finally, the fourth column vector (**fstatus**) denotes if a failure occurred (1=yes/0=no); for instance, the USP identified by id=1 does not experience any failure, while the USP identified by id=2 breaks down due to the c_2 cause of failure.

The results for the applied Cox PH and the Weibull models are shown in Tables 4 and 6 respectively. Note that we omit the subscript for the USP type without any loss of generality. The applied Cox PH model, and including the covariate "cause of failure," is defined as:

$$\log h(t) = \log h_0(t) + \sum_{k=1}^{K-1} \boldsymbol{\beta}_k \mathbf{x}_k, \quad (11)$$

where $h_0(t)$ is the baseline hazard, while \mathbf{x}_k is a dummy variable taking value one if the k -th cause of failure (c_k) has occurred, and zero otherwise. The cause c_4 is taken as a reference level. Therefore, there are three estimated coefficients $k = 1, 2, 3$.

To better explain the interpretation of the results in terms of HRs, in Table 5, we outline through an example the computation of the HRs estimates for the Cox PH model (Tab. 6). Therefore, in Tables 4 and 5, and by considering the reference level c_4 , we can observe that for probe-a, the failure risk is extremely high for the failure c_2 (132%). Instead, we have lower risks of failure for the causes c_1 and c_3 of 9.90% and 69.10% respectively. When considering probe-b, the pattern is different: i) the failure risk of c_1 with respect to c_1 is considerably low (lower than 98%); ii) the failure risk for c_1 is 86% approximately lower with respect to c_1 ; iii) the failure risk for cause c_2 is 71.40% higher. Moreover, all the estimated coefficients for both USPs are statistically significant, except for the failure c_1 (probe-a), with relatively low standard errors.

In what follows, the Weibull model expression, including the "cause of failure" as a covariate and related to a single observation, is reported:

$$\log \mathbf{T} = \boldsymbol{\beta}_0^* + \sum_{k=1}^{K-1} \boldsymbol{\beta}_k^* \mathbf{x}_k + \sigma\varepsilon. \quad (12)$$

where σ is the dispersion parameter to be estimated, ε is the vector of error term assumed to follow the Standard Extreme Value distribution, and \mathbf{x}_k is a dummy variable taking value one if cause of failure c_k has occurred and zero otherwise. The cause c_4 is taken as a reference level, therefore $k = 1, 2, 3$.

All the estimation details are illustrated in Table 6. When considering the estimated HRs for each cause of failure (reference level c_4), the calculated HRs are similar to those obtained through the Cox PH model (Tab. 4) for both probes, as expected. Therefore, there is a generally decreasing risk of failure due to the array-contact failure c_1 with respect to plastics-unknown failure c_4 .

Instead, for the probe-a and the Weibull model, the risk of failure is 10% lower for the array-contact failure c_1 with respect to the plastics-unknown one; while, when considering the probe-b, such hazard is 98% lower.

According to the chosen model, e.g. by discriminating between models (11) and (12), and by considering the specific parametrization and the distribution assumed for the time to failure, the computation of the HRs changes. In what follows, three examples of computation of the HRs are reported for the Cox PH, Weibull, and Exponential models; all the numerical examples are related to probe-a and the cause of failure c_1 versus c_4 .

1. Cox PH model, formula (11):

$$\text{HR}_k^C = \exp(\hat{\boldsymbol{\beta}}_k) \quad k = 1, 2, 3.$$

$$\text{HR}_1^C = \exp(\hat{\boldsymbol{\beta}}_1) = \exp(-0.1040) = 0.9010.$$

2. Weibull model, formula (12):

$$\text{HR}_k^W = \exp\left(\frac{-\hat{\boldsymbol{\beta}}_k^*}{\sigma}\right) \quad k = 1, 2, 3.$$

id	time(gg)	cause of failure	fstatus
1	455	1	0
1	455	2	0
1	455	3	0
1	455	4	0
2	410	1	0
2	410	2	1
2	410	3	0
2	410	4	0
.	.	.	.
.	.	.	.

Fig. 3. An example of an extended data-set for two statistical units.

$$HR_1^W = \exp\left(\frac{-\hat{\beta}_1^*}{\sigma}\right) = \exp\left(-\frac{0.0947}{0.9109}\right) = 0.9013.$$

3. Exponential model (formula (12) with $\sigma = 1/\gamma = 1$):

$$HR_k^E = \exp\left(-\hat{\beta}_k\right) \quad k = 1, 2, 3.$$

$$HR_1^E = \exp\left(-\hat{\beta}_1^*\right) = \exp(-0.0947) = 0.9096.$$

Furthermore, the main differences between the model results (Tabs. 4 and 6) lie in the estimated intercept and in specific estimated coefficients like the shape γ , or equivalently σ , equal to $1/\gamma$. As well known, the Weibull distribution changes its shape according to the γ values (see [16] for further details). The intercept represents the baseline situation; in this case-study it corresponds to the absence of failures; however, it is possible to define the intercept differently according to the specific situation. For example, a specific combination of types of products [26]. Furthermore, given the estimated intercept value, the corresponding HR represents the basic risk in the absence of the specific considered failure. For both USP types, it is equal to $2e-5$. The last remark is related to the estimated values of σ and γ . The estimated parameter $\hat{\gamma}$ is higher than one; this implies an increasing rate of failures with time, while the shape parameter $\hat{\sigma}$ is just a little lower than one. Therefore, the Weibull distribution, in this case, is very close to the Exponential one for both probes.

The statistical results shown in this Section are obtained by applying the LIFEREG and the PHREG procedures of the SAS software (version 9.4, Windows Platform). The LIFEREG procedure is used for the Weibull model estimation, while the PHREG for the Cox PH model. In general, the LIFEREG procedure supports also left and interval censoring, while the PHREG procedure supports right censoring only, (see [16] for

further comparative details). When considering the inclusion of random effects, the PHREG procedure allows for including only one random effect. This is not permitted in the LIFEREG procedure. However, for the application of several complex reliability models involving random effects, with different time distributions and different parametrizations, as well as several technical characterizations of the intercept term, the NLMIXED procedure could be particularly suitable [26]. Nevertheless, it must be noted that the NLMIXED procedure requires the computation of the likelihood function of the corresponding reliability model; for further details, refer to [16].

5 Discussion and final remarks

This paper focused on reliability statistical models in case of multiple causes of failures, evaluating several aims related to the modeling features. In particular, two distinct issues are faced: i) the differentiation between kinds of models, e.g. competing risks models and models for comparing causes of failure, by considering theoretical issues, also illustrated through a case-study on USPs; ii) within each approach, the specific parametrization and the HRs computation, as well as the model choice and interpretation are explored by highlighting empirical features.

The two different reliability modeling approaches illustrated in this manuscript are very suitable for companies in real production contexts, also allowing to treat fixed as well as random effects. In fact, random effects are particularly relevant because they can strongly affect the products' lifetime (for example, noises, and improvements). When considering competing risks models with random effects, the results obtained through the Cox PH and the Fine & Gray models are satisfactory. More precisely, the results show that probe-a has a better performance with respect to probe-b for all causes of failure. As to random effects, the presence of noises increases the risk of failure for probe-a by a higher amount compared to probe-b for all causes of failure. Furthermore, the presence of improvements decreases the rate of failure for probe-a by a higher amount compared to probe-b for causes c_1 and c_2 , while the opposite holds for cause c_3 . Considering cause c_4 , the risk reduction due to the presence of improvements does not differ for the two probes. When considering the Weibull and the Cox models for failure evaluation separately for each probe, the results are very satisfactory. Differently from the Cox model, in the Weibull model the baseline coefficients can be estimated and they are steady for both US probes. The failure risk due to cause c_2 is higher, while the failure risks due to causes c_1 and c_3 are lower, with c_4 as the reference level.

Nevertheless, further developments could be achieved through future researches aiming to improve data information, involving additional variables on ultrasonic probes in actual use, and also analyzing and comparing these model results with other reliability statistical models, in order to evaluate the accuracy of prediction, and the explanatory capability.

Table 4. Estimates and HRs by USP types for model (11).

Coef.	Probe-a Estimate	Std.Err.	<i>p</i> -value	HR	Probe-b Estimate	Std.Err.	<i>p</i> -value	HR
β_1	-0.1040	0.1614	0.5192	0.9010	-3.8918	1.0102	0.0001	0.0200
β_2	0.8420	0.1329	<0.0001	2.3210	0.5390	0.1798	0.0027	1.7140
β_3	-1.1750	0.2288	<0.0001	0.3090	-1.9459	0.4041	<0.0001	0.1430
β_4	0	-	-	1	0	-	-	1

Table 5. Interpretation of the HRs for model (11), by USP types.

Probe-a HR estimate	HR interpretation
0.9010	$(0.9010-1) \times 100 = -9.90\%$ → the risk of c_1 failure versus c_4 failure is -9.90% lower
2.3210	$(2.3210-1) \times 100 = 132\%$ → the risk of c_2 failure versus c_4 failure is 132% higher
0.3090	$(0.3090-1) \times 100 = -69.10\%$ → the risk of c_3 failure versus c_4 failure is -69.10% lower
Probe-b HR estimate	HR interpretation
0.0200	$(0.0200) \times 100 = -98.00\%$ → the risk of c_1 failure versus c_4 failure is -98% lower
1.7140	$(1.7140-1) \times 100 = 71.40\%$ → the risk of c_2 failure versus c_4 failure is 71.40% higher
0.1430	$(0.1430-1) \times 100 = -85.70\%$ → the risk of c_3 failure versus c_4 failure is -85.70% lower

Table 6. Estimates and HRs by USP types for model (12)¹.

Coef.	Probe-a Estimate	Std.Err.	<i>p</i> -value	HR	Probe-b Estimate	Std.Err.	<i>p</i> -value	HR
β_0	9.7148	0.2123	<0.0001	2e-5	10.2025	0.3680	<0.0001	2e-5
β_1^*	0.0947	0.1471	0.5196	0.9013	3.6383	0.9924	0.0002	0.0204
β_2^*	-0.7670	0.1274	<0.0001	2.3211	-0.5039	0.1733	0.0036	1.7143
β_3^*	1.0709	0.2156	<0.0001	0.3086	1.8192	0.4074	<0.0001	0.1429
β_4^*	0	-	-	1	0	-	-	1
σ	0.9109	0.0470	-	-	0.9349	0.0784	-	-
γ	1.0978	0.0567	-	-	1.0697	0.0897	-	-

¹ It must be noted that, according to the usual and consolidated notation for the Weibull distribution, the used symbols σ and γ have a different meaning in comparison with that reported in Section 3.3.

Funding

This research received no external funding.

Conflicts of interest

The authors have nothing to disclose any kind of conflict of interest.

Data availability statement

The authors may provide a similar dataset with counterfeit data. Unfortunately, the original data are confidential, and therefore, they cannot be shared.

Author contribution statement

Conceptualization, Rossella Berni, Francesco Bertocci; Methodology, Alessandro Magrini, Nedka D. Nikiforova; Validation, Francesco Bertocci; Formal Analysis, Alessandro Magrini, Nedka D. Nikiforova; Data Curation, Francesco Bertocci, Alessandro Magrini; Writing – Original Draft Preparation, Alessandro Magrini, Nedka D. Nikiforova; Writing – Review & Editing, Rossella Berni.

References

1. R Core Team. *R: A language and environment for S statistical computing*. (R S Foundation for Statistical Computing, Vienna, AT, 2022)
2. T.M. Therneau, Coxme: mixed effects Cox models. R package version 2.2-18.1; 2022. <https://CRAN.R-project.org/package=coxme>
3. D.R. Cox, Regression models and life tables (with discussion), *J. Royal Stat. Soc. Ser. B.* **34**, 187–220 (1972)
4. J.P. Fine, R.J. Gray, A proportional hazards model for the subdistribution of a competing risk, *J. Am. Stat. Assoc.* **94**, 496–509 (1999)
5. F. Bertocci, A. Grandoni, A.T. Djuric-Rissner, Scanning acoustic microscopy (SAM): a robust method for defect detection during the manufacturing process of ultrasound probes for medical imaging, *Sensors* **19**, 4868–4886 (2019)
6. F. Bertocci, A. Grandoni, M. Fidanza, R. Berni, A guideline for implementing a robust optimization of a complex multi-stage manufacturing process, *Appl. Sci.* **11**, 1–19 (2021)
7. A.M. Vitikainen, J.I. Peltonen, E. Vartiainen, Routine ultrasound quality assurance in a multi-unit radiology department: a retrospective evaluation of transducer failures, *Ultrasound Med. Biol.* **43**, 1930–1937 (2017)
8. W. Ding, M. Bavencoffe, M. Lethiecq, Modeling and experimental characterization of bonding delaminations in single-element ultrasonic transducer, *Materials* **14**, 2269 (2021)
9. N.J. Dudley, D.J. Woolley, Blinded comparison between an in-air reverberation method and an electronic probe tester in the detection of ultrasound probe faults, *Ultrasound Med. Biol.* **43**, 2954–2958 (2017)
10. J. Vachutka, L. Dolezal, C. Kollmann, J. Klein, The effect of dead elements on the accuracy of Doppler ultrasound measurements, *Ultrason Imaging.* **36**, 18–34 (2014)
11. R. Lorentsson, N. Hosseini, L.G. Mansson, M. Bath, Evaluation of an automatic method for detection of defects in linear and curvilinear ultrasound transducers, *Phys. Med.* **84**, 33–40 (2021)
12. R. Lorentsson, N. Hosseini, J.O. Johansson, W. Rosenberg, B. Stenborg, L.G. Mansson, M. Bath, Method for automatic detection of defective ultrasound linear array transducers based on uniformity assessment of clinical images – a case study, *J. Appl. Clin. Med. Phys.* **19**, 265–274 (2018)
13. L. Wang, B. Li, B. Hu, G. Shen, Y. Zhen, Y. Zheng, Failure mode effect and criticality analysis of ultrasound device by classification tracking, *BMC Health Serv. Res.* **22**, 1–10 (2022)
14. E. Sassaroli, A. Scorza, C. Crake, S.A. Sciuto, M. Park, Breast ultrasound technology and performance evaluation of ultrasound equipment: B- Mode. *IEEE Trans. Ultrason. Ferroelectr. Fr eq. Control.* **64**, 192–205 (2017)
15. W.Q. Meeker, L.A. Escobar. *statistical methods for reliability data* (John Wiley & Sons, New York, 1998)
16. P.D. Allison: *Survival analysis using SAS – a practical guide*, 2nd edn. (SAS Institute, Cary (US-NC), 2010)
17. L. Duchateau, P. Janssen, *The frailty model* (Springer, New York, 2007)
18. A. Wienke, *Frailty models in survival analysis* (CRC Press, Amsterdam, The Netherlands, 2010)
19. J.W. Vaupel, K.G. Manton, E. Stallard, The impact of heterogeneity in individual frailty on the dynamics of mortality, *Demography* **16**, 439–454 (1979)
20. T.A. Balan, M. Putter, A tutorial on frailty models, *Stat. Methods Med. Res.* **29**, 3424–3454 (2020)
21. J.H. Abbring, G.J. Van Den Berg, The unobserved heterogeneity distribution in duration analysis, *Biometrika* **94**, 87–99 (2007)
22. S. Ripatti, J. Palmgren, Estimation of multivariate frailty models using penalized partial likelihood, *Biometrics* **56**, 1016–1022 (2000)
23. T.M. Therneau, P.M. Grambsch, V.S. Pankratz, Penalized survival models and frailty, *J. Computat. Graph. Stat.* **12**, 156–175 (2003)
24. C. Elbers, G. Ridder, True and spurious duration dependence: the identifiability of the proportional hazard model, *Rev. Economic. Studies.* **49**, 403–409 (1982)
25. Y. Lee Y, J.A.Nelder, Hierarchical generalized linear models (with discussion), *J. R. Stat. Soc. B* **58**, 619–678 (1996)
26. R. Berni, M. Catelani, C. Fiesoli, V.L. Scarano, A comparison of alloy-surface finish combinations considering different component package types and their impact on soldering reliability, *IEEE Trans. Reliab.* **65**, 272–281 (2016)

Cite this article as: Rossella Berni, Francesco Bertocci, Alessandro Magrini, Nedka D. Nikiforova, Reliability with multiple causes of failures: Modeling and practice through a case study on ultrasound probes for medical imaging, *Int. J. Metrol. Qual. Eng.* **15**, 19 (2024)