Received: 1 October 2020 | Revised: 23 January 2021 | Accepted: 28 February 2021

DOI: 10.1002/sim.8961

RESEARCH ARTICLE

Statistics in Medicine WILEY

Degrees of necessity and of sufficiency: Further results and extensions, with an application to covid-19 mortality in Austria

Andreas Gleiss[1] | **Robin Henderson**[2] | **Michael Schemper**[1]

[1]Section for Clinical Biometrics, Center for Medical Statistics, Informatics, and Intelligent Systems, Medical University of Vienna, Vienna, Austria

[2]School of Mathematics, Statistics and Physics, Newcastle University, Newcastle upon Tyne, UK

Correspondence
Andreas Gleiss, Section for Clinical Biometrics, Center for Medical Statistics, Informatics, and Intelligent Systems, Medical University of Vienna, Vienna, Austria.
Email: andreas.gleiss@meduniwien.ac.at

The purpose of this paper is to extend to ordinal and nominal outcomes the measures of degree of necessity and of sufficiency defined by the authors for dichotomous and survival outcomes in a previous paper. A cause, represented by certain values of prognostic factors, is considered necessary for an event if, without the cause, the event cannot develop. It is considered sufficient for an event if the event is unavoidable in the presence of the cause. The degrees of necessity and sufficiency, ranging from zero to one, are simple, intuitive functions of unconditional and conditional probabilities of an event such as disease or death. These probabilities often will be derived from logistic regression models; the measures, however, do not require any particular model. In addition, we study in detail the relationship between the proposed measures and the related explained variation summary for dichotomous outcomes, which are the common root for the developments for ordinal, nominal, and survival outcomes. We introduce and analyze the Austrian covid-19 data, with the aim of quantifying effects of age and other potentially prognostic factors on covid-19 mortality. This is achieved by standard regression methods but also in terms of the newly proposed measures. It is shown how they complement the toolbox of prognostic factor studies, in particular when comparing the importance of prognostic factors of different types. While the full model's degree of necessity is extremely high (0.933), its low degree of sufficiency (0.179) is responsible for the low proportion of explained variation (0.193).

KEYWORDS
explained variation, logistic regression, necessary condition, ordinal outcomes, sufficient condition

1 | INTRODUCTION

The notions of sufficient and of necessary conditions are common in formal logic and mathematics, and in empirical sciences such as medicine and epidemiology as well. For the latter, "conditions" are thought of as causes leading to outcomes, often events like diseases or death. A cause is considered necessary for an event if, without the cause, the event

TABLE 1 Covid-19 disease after infection by SARS-CoV-2 in Austria (as of July 22, 2020)

Age group	Number SARS-CoV-2 positive	Number (%) of deaths due to covid-19	Number (%) of late[a] deaths due to covid-19
Below 65	13 323	38 (0.3%)	19 (0.1%)
65 or older	3712	656 (17.7%)	415 (11.2%)
Total	17 035	694 (4.1%)	434 (2.5%)

Note: Patients with potential follow-up of at least 6 weeks.
[a] >1 week after diagnosis of infection.

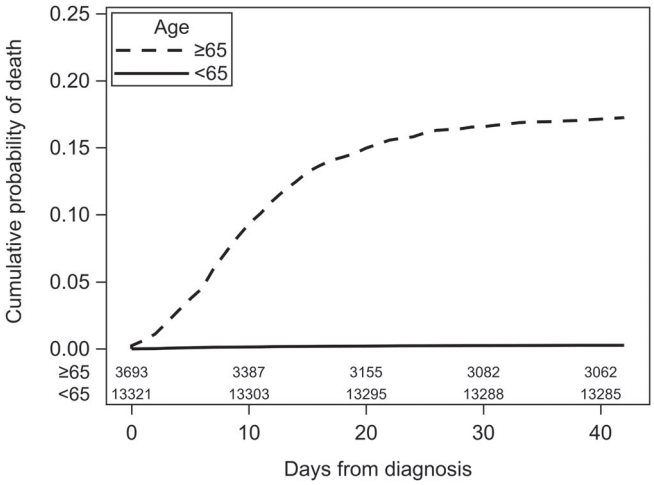

FIGURE 1 Kaplan–Meier estimates of survival from diagnosis to death (censored at 6 weeks after diagnosis). Numbers at risk given above x-axis

cannot develop. It is considered sufficient for an event if the event is unavoidable in the presence of the cause. In mathematics, sufficient and necessary conditions either fully apply ("100%"), or not ("0%"). In empirical sciences, the effects of causes on events only exceptionally achieve the extremes of 0% or of 100%, the latter, for example, for fully deterministic relationships. Thus, it is of interest to quantify the *degrees* to which certain levels of prognostic factors (ie, potential causes) are necessary or sufficient for events to occur.

In an earlier paper, we have introduced measures *DN* and *DS* for dichotomous outcomes which suitably quantify the degrees of necessity and of sufficiency, respectively.[1] Before reviewing them we motivate their use by means of data of the covid-19 pandemic in Austria summarized by Table 1.[2] We are interested in the effect of age (using an internationally confirmed cutpoint of 65)[3-5] on the survival of infected individuals.

This effect turns out to be very strong, in terms of relative risk it is $RR = 62.1$, however, the proportion of variation in survival status attributable to age is relatively low: $EV = 0.132$. The explained variation (EV) quantifies on a 0-1 scale the degree to which the outcome is determined by a prognostic factor.[6] EV (=0.132) can be decomposed into the product of DN (=0.930) and DS (=0.142).[1] Thus the Austrian data confirm that age is quite necessary for a covid-19 death (most of the deaths occurred in the age group 65 and beyond) but, fortunately, it is not sufficient for death as most of the elder age group survived after infection. This is a simple illustrative summary of the data we will return to in Section 6.

Furthermore, the covid-19 data also contain individual information on the time from diagnosis (confirmed positive test result) until covid-19 related death. Because of rather complete follow-up of covid-19 cases we decided to censor all times by 6 weeks (a time frame of supposed medical interest), and assumed all individuals without recorded death within this period to be alive at 6 weeks after infection. Estimates of cumulative incidences by age groups are shown in Figure 1, the corresponding hazard ratio, according to Cox regression, being $HR = 69.7$. We obtain $DN_1^{surv} = 0.930$ and $DS_1^{surv} = 0.139$ for the corresponding DN and DS measures for survival times,[1] and $V = 0.129$, the corresponding variation in survival, explained by age group.[7] As with the dichotomous viewpoint, age ≥65 is almost definitely required to die, but the outlook for these infected individuals is not fully daunting.

In Section 2 we provide a recap of the DN and DS measures proposed previously.[1] In Section 3 we extend to ordinal and in Section 4 to nominal outcomes. In Section 5 we examine the three-way relationships between DN, DS and EV, concentrating for simplicity on dichotomous outcomes. Section 6 gives a full analysis of the Austrian covid-19 data along

with the values of the proposed measures with dichotomous outcomes, first introduced in our earlier paper in this journal, and with ordinal and nominal outcomes introduced in Sections 3 and 4 of this paper.

2 | RECAP OF THE MEASURES DN AND DS

The measures, quantifying the degrees of necessity and of sufficiency, *DN* and *DS*, respectively, and ranging from 0 to 1, have been defined as follows:[1]

For a dichotomous outcome let D denote the event of interest (eg, death) with probability $P(D) \neq 0, 1$. For a dichotomous or continuous prognostic factor X we propose to define *DN* and *DS* in a population as

$$DN_1 = \sqrt{E_< \left(\frac{P(D) - P(D|X)}{P(D) - 0} \right)^2} \tag{1}$$

and

$$DS_1 = \sqrt{E_> \left(\frac{P(D|X) - P(D)}{1 - P(D)} \right)^2} \tag{2}$$

where $E_<$ and $E_>$ denote expectation conditional on $\{X: P(D|X) < P(D)\}$ and on $\{X: P(D|X) > P(D)\}$, respectively. Note that X might also represent a predictor from several prognostic factors. Alternatively, we define

$$DN_2 = E_< \left(\frac{P(D) - P(D|X)}{P(D) - 0} \right) \tag{3}$$

and

$$DS_2 = E_> \left(\frac{P(D|X) - P(D)}{1 - P(D)} \right) \tag{4}$$

For an unfavorable outcome event D (like death) the range of values for X with $P(D|X) < P(D)$ defines its *protective* levels in the sense that for these levels the conditional death probability falls below the unconditional one. For a favorable outcome the definition for *DN* above still applies but the same range defines *harmful* levels of X. It is therefore important to clearly denote the level of the outcome to which *DN* and *DS* measures refer. In the remaining text we assume an unfavorable outcome, if not explicitly stated otherwise, and designate levels of X with $P(D|X) < P(D)$ as protective and with $P(D|X) > P(D)$ as harmful.

By explicitly using 0 in the denominator of (1) we emphasized that the actual departure of $P(D|X)$ from $P(D)$ is standardized to the maximal (hypothetical) departure when $P(D|X)$ assumes a value of 0. Likewise, the actual departure of $P(D|X)$ from $P(D)$ is standardized by the denominator of (2), to the maximal (hypothetical) departure when $P(D|X)$ assumes a value of 1. Thus, *DN* and *DS* achieve their maximal values of 1 if conditional death probabilities are at the extreme values of 0 and 1, respectively. They share the same structure and are independently sensitive to relevant characteristics of a data set. All pairs of values (*DN, DS*) are possible within $[0, 1]^2$. The measures are defined independent of the model or prediction tool used to produce conditional probabilities. *DN* and *DS* have been demonstrated to provide an approximate, in some cases an exact multiplicative decomposition of explained variation,[1] as defined by Schemper for dichotomous outcomes.[6] Furthermore, *DN* implicitly generalizes the established *attributable fraction* or *risk* for dichotomous prognostic factors and dichotomous outcomes,[8] to continuous prognostic factors. In a setting with multiple prognostic factors marginal and partial results of *DN* and *DS* are provided akin to marginal and partial odds ratios from multiple logistic regression.

We refer the reader interested in more detail to our earlier paper in this journal.[1]

3 | DN AND DS FOR ORDINAL OUTCOMES

In this and the following section we present generalizations of *DN* and *DS* to ordinal and nominal outcomes. We will restrict the presentation to variant 1 but note that analogous generalizations are possible for variant 2.

For an ordinal outcome Y with categories $k = 0, 1, \ldots, K$ let us assume, without loss of generality, that 0 represents the "best" and K the "worst" outcome category. Such an outcome might be intrinsically ordinal like, for example, the ECOG performance status,[9] ranging from 0 (fully active) to 5 (dead), or derived by categorization from a continuous or survival scale. Compared to the case with dichotomous outcome, where unconditional and conditional probabilities, $P(D)$ and $P(D|X)$, of the (unfavorable) event of interest, D, are contrasted, we now consider $P(Y \geq k)$ and $P(Y \geq k|X)$ for all $k = 1, \ldots, K$.

With $I(.)$ denoting the indicator function, let $D_{k+} = I(Y \geq k)$ for $k = 1, \ldots, K$. For each of the resulting K ordered dichotomizations D_{k+} of the $K+1$ levels of the ordinal outcome we define

$$DN_{1,k+} = \sqrt{\mathrm{E}_{<}\left(\frac{\mathrm{P}(D_{k+}) - \mathrm{P}(D_{k+}|X)}{\mathrm{P}(D_{k+})}\right)^2} \tag{5}$$

$$DS_{1,k+} = \sqrt{\mathrm{E}_{>}\left(\frac{\mathrm{P}(D_{k+}|X) - \mathrm{P}(D_{k+})}{1 - \mathrm{P}(D_{k+})}\right)^2} \tag{6}$$

using (1) and (2), and $\mathrm{E}_{<}(.)$ and $\mathrm{E}_{>}(.)$ denoting conditional expectations over $\{X : \mathrm{P}(D_{k+}|X) < \mathrm{P}(D_{k+})\}$ and $\{X : \mathrm{P}(D_{k+}|X) > \mathrm{P}(D_{k+})\}$, respectively. DN_1^{ord} and DS_1^{ord} for ordinal outcomes can then be calculated as weighted averages of $DN_{1,k+}$ and $DS_{1,k+}$, respectively:

$$DN_1^{ord} = \frac{1}{\mathrm{P}(Y > 0)} \sum_{k=1}^{K} DN_{1,k+} \mathrm{P}(Y = k) \tag{7}$$

$$DS_1^{ord} = \frac{1}{\mathrm{P}(Y > 0)} \sum_{k=1}^{K} DS_{1,k+} \mathrm{P}(Y = k) \tag{8}$$

Note that the weighted sums in (7) and (8) do not include $P(Y = 0)$ in agreement with definitions (1) and (2) for DN_1 and DS_1, respectively, to which DN_1^{ord} and DS_1^{ord} specialize for dichotomous outcomes. The averaging on the original scale of the $DN_{1,k+}$ is motivated by the analogous approach used for survival outcomes (see Appendix A). Simulations indicate that averaging on a quadratic scale would result in higher median absolute discrepancies between EV and $DN_1^{ord} \cdot DS_1^{ord}$ (results not shown).

As with dichotomous outcomes it is possible and useful to calculate *partial* degrees of necessity and of sufficiency. We propose to calculate the partial DN_1^{ord} as the difference of DN_1^{ord} values with and without the factor of interest, since the weights $\mathrm{P}(Y = k)$ are the same in the model containing and the model not containing this factor.

For estimation of DN_1^{ord} and DS_1^{ord} we replace population values in (5) to (8) by estimates from a representative sample. For a sample of n individuals let $\hat{p}_k$ denote the unconditional probability for the kth outcome category, $\hat{S}(k) = \sum_{j=k}^{K} \hat{p}_j$, the unconditional estimate of $\mathrm{P}(Y \geq k)$, and $\hat{S}(k|x_i)$ the corresponding conditional estimate for the ith individual $(1 \leq i \leq n)$ which could be obtained from a (multiple) ordinal logistic regression using covariate vector x_i.[10]

We arrive at

$$\widehat{DN}_1^{ord} = \frac{1}{\hat{S}(1)} \sum_{k=1}^{K} \hat{p}_k \left[\frac{1}{n_{<}(k)} \sum_{i=1}^{n} I(\hat{S}(k|x_i) < \hat{S}(k)) \left(\frac{\hat{S}(k) - \hat{S}(k|x_i)}{\hat{S}(k)}\right)^2\right]^{1/2} \tag{9}$$

and

$$\widehat{DS}_1^{ord} = \frac{1}{\hat{S}(1)} \sum_{k=1}^{K} \hat{p}_k \left[\frac{1}{n_{>}(k)} \sum_{i=1}^{n} I(\hat{S}(k|x_i) > \hat{S}(k)) \left(\frac{\hat{S}(k|x_i) - \hat{S}(k)}{1 - \hat{S}(k)}\right)^2\right]^{1/2} \tag{10}$$

with $n_{<}(k)$ and $n_{>}(k)$ denoting the number of subjects i with $\hat{S}(k|x_i) < \hat{S}(k)$ and $\hat{S}(k|x_i) > \hat{S}(k)$, respectively.

Though it is obvious that the case with ordinal outcomes should at least be similar to the case of survival outcomes without censoring,[1] this is worked out in detail in Appendix A.

In the same way, the explained variation measure for survival outcomes, V_W,[7] simplifies in uncensored samples to the appropriate measure for ordinal outcomes, EV^{ord}:

$$EV^{ord} = \frac{1}{\mathrm{P}(Y>0)} \sum_{k=1}^{K} EV_{k+} \mathrm{P}(Y=k) \tag{11}$$

where

$$EV_{k+} = 1 - \frac{\mathrm{E}[\mathrm{P}(D_{k+}|X)(1-\mathrm{P}(D_{k+}|X))]}{\mathrm{P}(D_{k+})(1-\mathrm{P}(D_{k+}))} \tag{12}$$

which is analogous to the definition for dichotomous outcomes.[6]

We shall illustrate the approach for ordinal outcomes by means of the covid-19 example in Section 6.

4 | DN AND DS FOR NOMINAL OUTCOMES

Similar to our approach for ordinal outcomes in the previous section also our proposals for nominal outcomes are based on certain dichotomizations of the considered outcome $Y \in \{0,1,\dots,K\}$.

If there is no reference category among the $K+1$ categories, then each outcome category can be separately contrasted to all remaining categories, that is, k vs $\neg\, k$ for all $k = 0,\dots,K$. By defining $D_k = I(Y=k)$ for $k = 0,\dots,K$ let

$$DN_{1,k} = \sqrt{\mathrm{E}_{<}\left(\frac{\mathrm{P}(D_k) - \mathrm{P}(D_k|X)}{\mathrm{P}(D_k)}\right)^2} \tag{13}$$

$$DS_{1,k} = \sqrt{\mathrm{E}_{>}\left(\frac{\mathrm{P}(D_k|X) - \mathrm{P}(D_k)}{1-\mathrm{P}(D_k)}\right)^2} \tag{14}$$

However, we refrain from proposing a single measure given as a (weighted) average of the $DN_{1,k}$ (or $DS_{1,k}$) since the $DN_{1,k}$ (or $DS_{1,k}$) neutralize each other across categories k. This is easily seen for the dichotomous case $K=1$ where $DN_{1,0} = DS_{1,1}$ and $DS_{1,0} = DN_{1,1}$. Depending on the weights $\mathrm{P}(Y=k)$ the weighted result might be misleading. The inspection and comparison of the individual $DN_{1,k}$ and $DS_{1,k}$ for $k=0,\dots,K$ is favored instead.

If there is a unique reference category, say category 0, then each of the remaining (unordered) outcome categories can be contrasted separately to this reference, that is, k vs 0 for all $k = 1,\dots,K$. By defining $D'_k = I(Y=k \mid Y \in \{0,k\})$ for $k = 1,\dots,K$ let

$$DN'_{1,k} = \sqrt{\mathrm{E}_{<}\left(\frac{\mathrm{P}(D'_k) - \mathrm{P}(D'_k|X)}{\mathrm{P}(D'_k)}\right)^2} \tag{15}$$

$$DS'_{1,k} = \sqrt{\mathrm{E}_{>}\left(\frac{\mathrm{P}(D'_k|X) - \mathrm{P}(D'_k)}{1-\mathrm{P}(D'_k)}\right)^2} \tag{16}$$

DN_1^{nom0} and DS_1^{nom0} for nominal outcomes with reference 0 are then given as weighted averages of $DN'_{1,k}$ and $DS'_{1,k}$, respectively:

$$DN_1^{nom0} = \frac{1}{\mathrm{P}(Y>0)} \sum_{k=1}^{K} DN'_{1,k} \mathrm{P}(Y=k) \tag{17}$$

$$DS_1^{nom0} = \frac{1}{\mathrm{P}(Y>0)} \sum_{k=1}^{K} DS'_{1,k} \mathrm{P}(Y=k) \tag{18}$$

Estimation of conditional probabilities of D'_k might be based on a generalized logit model (eg, by using the link = glogit option in SAS proc logistic).

The EV measure for nominal outcomes with reference category 0, EV^{nom0}, is structurally identical to EV^{ord} as defined by (11) and (12):

$$EV^{nom0} = \frac{1}{\mathrm{P}(Y>0)} \sum_{k=1}^{K} EV'_k \mathrm{P}(Y=k) \tag{19}$$

where

$$EV'_k = 1 - \frac{\mathrm{E}[\mathrm{P}(D'_k|X)(1-\mathrm{P}(D'_k|X))]}{\mathrm{P}(D'_k)(1-\mathrm{P}(D'_k))} \tag{20}$$

We shall illustrate the approach for nominal outcomes by means of the covid-19 example in Section 6.

5 | DN, DS, AND EV FOR A DICHOTOMOUS OUTCOME

This section serves two purposes. First, the two variants of *DN* and *DS*, as given in Equations (1) and (2) on the one hand, and in Equations (3) and (4) on the other hand, will be contrasted with each other. Second, based on these results we will be able to gain a deeper understanding of the connection of *DN* and *DS* with *EV* for dichotomous outcomes.

5.1 | The two variants of DN and DS

We first investigate the decomposition of *EV* into the variation (ie, difference) between predictions in the protective and in the harmful range of a prognostic factor, and the variation of predictions within each of these ranges. In the following, a similar decomposition for *DN* and *DS* will enable us to show that DN_1 and DS_1 comprise both, the "within" and the "between" component, while DN_2 and DS_2 reflect only the latter.

For enhanced readability, we denote the conditional probability, expectation, and variance in the protective range of X by $\mathrm{P}_<(.)$, $\mathrm{E}_<(.)$, and $\mathrm{Var}_<(.)$, respectively, for example, $\mathrm{P}_<(D) = \mathrm{P}_X(D|\,\mathrm{P}(D|X) < \mathrm{P}(D))$. A similar notation is used for these entities in the harmful range of X by $\mathrm{P}_>(.)$, $\mathrm{E}_>(.)$, and $\mathrm{Var}_>(.)$.

An explained variation measure has been defined as[6]

$$EV = 1 - \frac{\mathrm{E}[\mathrm{P}(D|X)(1-\mathrm{P}(D|X))]}{\mathrm{P}(D)(1-\mathrm{P}(D))} \tag{21}$$

which is equivalent to

$$EV = \frac{\mathrm{E}(\mathrm{P}(D|X)-\mathrm{P}(D))^2}{\mathrm{P}(D)(1-\mathrm{P}(D))} \tag{22}$$

if global calibration of the "true" model, that is, $\mathrm{E}(\mathrm{P}(D|X)) = \mathrm{P}(D)$, is assumed.[1] The numerator in Equation (22) is then equal to the variance of $\mathrm{P}(D|X)$. In addition, we assume that the "true" model is also locally calibrated in the protective and in the harmful range of X such that $\mathrm{E}_<(\mathrm{P}(D|X)) = \mathrm{P}_<(D)$ and $\mathrm{E}_>(\mathrm{P}(D|X)) = \mathrm{P}_>(D)$. With $\alpha = \mathrm{P}_X(\mathrm{P}(D|X) > \mathrm{P}(D))$ and noting that

$$\mathrm{P}(D) = (1-\alpha)\mathrm{P}_<(D) + \alpha \mathrm{P}_>(D) \tag{23}$$

we arrive at the following decomposition for the numerator of *EV* (see Appendix B):

$$\begin{aligned}\mathrm{Var}(\mathrm{P}(D|X)) = {} & (1-\alpha)\mathrm{Var}_<(\mathrm{P}(D|X)) + \alpha \mathrm{Var}_>(\mathrm{P}(D|X)) \\ & + \alpha(1-\alpha)[\mathrm{P}_>(D) - \mathrm{P}_<(D)]^2\end{aligned} \tag{24}$$

The first two terms of the right hand side of (24) are a weighted average of the variations of the conditional predictions within the protective and the harmful ranges around their respective conditional expectations. The third term measures

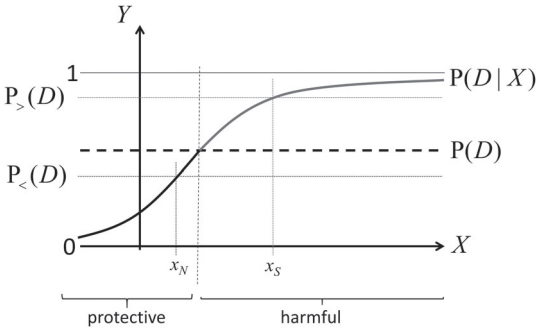

FIGURE 2 Schematic representation of unconditional probability P(D) and conditional probabilities P($D|X$), P$_<$(D) and P$_>$(D). For x_S and x_N see text

the variation between the protective and the harmful range by means of the same conditional expectations. For scalar X, Figure 2 illustrates the unconditional and various conditional probabilities used in this decomposition.

Similarly, noting that (23) implies $\mathrm{P}(D) - \mathrm{P}_<(D) = \alpha(\mathrm{P}_>(D) - \mathrm{P}_<(D))$ the numerator of ${DN_1}^2$ from (1) can be decomposed into

$$\mathrm{E}_<(\mathrm{P}(D) - \mathrm{P}(D|X))^2 = \mathrm{Var}_<(\mathrm{P}(D|X)) + \alpha^2[\mathrm{P}_>(D) - \mathrm{P}_<(D)]^2 \tag{25}$$

while that of ${DS_1}^2$ from (2) is equal to

$$\mathrm{E}_>(\mathrm{P}(D|X) - \mathrm{P}(D))^2 = \mathrm{Var}_>(\mathrm{P}(D|X)) + (1-\alpha)^2[\mathrm{P}_>(D) - \mathrm{P}_<(D)]^2 \tag{26}$$

Thus, in variant 1 the degree of necessity comprises the variation of conditional probability within the protective range plus a proportion of the "between" variation, and analogously for the degree of sufficiency and the harmful range.

In contrast to variant 1, the numerators of DN_2 and of DS_2 cover only the "between" part:

$$\mathrm{E}_<(\mathrm{P}(D) - \mathrm{P}(D|X)) = \alpha[\mathrm{P}_>(D) - \mathrm{P}_<(D)] \tag{27}$$

and

$$\mathrm{E}_>(\mathrm{P}(D|X) - \mathrm{P}(D)) = (1-\alpha)[\mathrm{P}_>(D) - \mathrm{P}_<(D)] \tag{28}$$

resulting in

$$DN_2 = \frac{\alpha}{\mathrm{P}(D)}[\mathrm{P}_>(D) - \mathrm{P}_<(D)] \tag{29}$$

and

$$DS_2 = \frac{1-\alpha}{1-\mathrm{P}(D)}[\mathrm{P}_>(D) - \mathrm{P}_<(D)] \tag{30}$$

Thus, for obtaining high degrees of necessity or sufficiency in variant 2, a small "between" variation $[\mathrm{P}_>(D) - \mathrm{P}_<(D)]$ can only be compensated by a large harmful range for DN_2, that is, large α in relation to P(D), or a large protective range for DS_2.

However, this simple structure allows to interpret DN_2 from (3) as the attributable risk at a value x_N that results in $\mathrm{P}(D|X = x_N) = \mathrm{P}_<(D)$:

$$DN_2 = \frac{\mathrm{P}(D) - \mathrm{P}_<(D)}{\mathrm{P}(D) - 0} = \frac{\mathrm{P}(D) - \mathrm{P}(D|X = x_N)}{\mathrm{P}(D) - 0} \tag{31}$$

For a monotone model P($D|X$), x_N is unique (see Figure 2). Analogously,

$$DS_2 = \frac{\mathrm{P}(D|X = x_S) - \mathrm{P}(D)}{1 - \mathrm{P}(D)} \tag{32}$$

with x_S such that $\mathrm{P}(D|X = x_S) = \mathrm{P}_>(D)$, is interpretable as *reverse attributable risk*.[1]

5.2 | Discrepancy between EV and the product of DN and DS

In our previous paper we have shown that $EV = DN_1 \cdot DS_1 = DN_2 \cdot DS_2$ if a single dichotomous prognostic factor is used for conditional prediction.[1] In more general settings, however, discrepancies between EV and $DN{\cdot}DS$ were observed with positive or negative sign, which, for real data applications, proved to be reasonably small. This elegant multiplicative decomposition is conceptually interesting, and we now can provide further detail for the case of dichotomous outcomes and continuous prognostic factors. This subsection presents theoretical and empirical results on the deviation from this decomposition. As definitions of DN and DS for possibly censored survival outcomes and for ordinal and nominal outcomes build on DN and DS for dichotomous outcomes, the conclusions from this section extend to these more complex outcomes as well if the degrees of necessity and of sufficiency are homogenous across time or categories.

First, we are interested in $\Delta_1 = EV - DN_1 \cdot DS_1$. From Equations (1), (2), (25), and (26) we conclude:

$$DN_1 = \sqrt{\mathrm{Var}_<(\mathrm{P}(D\,|\,X)) + \alpha^2[\mathrm{P}_>(D) - \mathrm{P}_<(D)]^2}/P(D) \tag{33}$$

$$DS_1 = \sqrt{\mathrm{Var}_>(\mathrm{P}(D\,|\,X)) + (1-\alpha)^2[\mathrm{P}_>(D) - \mathrm{P}_<(D)]^2}/(1 - P(D)) \tag{34}$$

Equations (22) and (24) result in

$$\begin{aligned} EV &= \mathrm{Var}(P(D\,|\,X))/[\mathrm{P}(D)(1-\mathrm{P}(D))] \\ &= \frac{(1-\alpha)\mathrm{Var}_<(\mathrm{P}(D\,|\,X)) + \alpha\mathrm{Var}_>(\mathrm{P}(D\,|\,X)) + \alpha(1-\alpha)[\mathrm{P}_>(D) - \mathrm{P}_<(D)]^2}{\mathrm{P}(D)(1-\mathrm{P}(D))} \end{aligned} \tag{35}$$

There is no simplified general expression for Δ_1; for properties of Δ_1 under special conditions see Appendix C.

Second, using Equations (27) and (28), EV can be decomposed into $DN_2 \cdot DS_2$ plus a non-negative term comprising the "within" components:

$$\Delta_2 = EV - DN_2 \cdot DS_2 = \frac{(1-\alpha)\mathrm{Var}_<(\mathrm{P}(D\,|\,X)) + \alpha\mathrm{Var}_>(\mathrm{P}(D\,|\,X))}{\mathrm{P}(D)(1-\mathrm{P}(D))} \geq 0 \tag{36}$$

Thus, the larger the variation of conditional probabilities for D within the protective and harmful levels of X, the larger is Δ_2. The identity of EV and $DN_2 \cdot DS_2$ is exact for dichotomous X ($\mathrm{Var}_< = \mathrm{Var}_> = 0$), though it approaches zero for situations with extremely strong covariate effects and all probabilities stacked near zero or one. Furthermore, since $DN_1 \geq DN_2$ and $DS_1 \geq DS_2$ it follows that $\Delta_1 \leq \Delta_2$, which is consistent with the fact that variant 2 does not contain any "within" component.

In order to assess the magnitudes of Δ_1 and Δ_2, we assume the conditional predictions to come from a simple logistic regression model

$$\log\left(\frac{\mathrm{P}(D\,|\,X)}{1-\mathrm{P}(D\,|\,X)}\right) = \beta_0 + \beta_1 X \tag{37}$$

with a single marginally standard normally distributed prognostic factor X. The regression coefficient β_0 is set to 0, ±1, ±2, ±3, ±4, while β_1 varies between -2 and $+2$ in steps of 0.1. As a result, $\mathrm{P}(D)$ equals 0.5 for $\beta_0 = 0$, while it ranges from 0.65 to 0.73 for $\beta_0 = 1$, and from 0.87 to 0.95 for $\beta_0 = 3$, the maximum being obtained at $\beta_1 = 0$, the minimum at $\beta_1 = \pm 2$. All (conditional) expectations involved were numerically approximated using the *integrate*() function in R version 3.5.1. The dependence of Δ_1 and Δ_2 on β_0 and the odds ratio,$OR = \exp(\beta_1)$, is visualized in Figure 3.

For model (37) we conclude from Figure 3:

- The discrepancy between EV and the product of DN and DS is much smaller for variant 1.
- If the standardized odds ratio, $OR = \exp(\beta_1)$, equals 1 then $\Delta_1 = \Delta_2 = 0$. In this case $EV = DN_1 = DS_1 = DN_2 = DS_2 = 0$ holds, which is a required property for these measures.[1]
- $\Delta_1 \geq 0$ and $\Delta_2 \geq 0$

Although not apparent for the odds ratios shown, both Δ_1 and Δ_2 peak and then decline towards zero as either OR or $1/OR$ increase indefinitely.

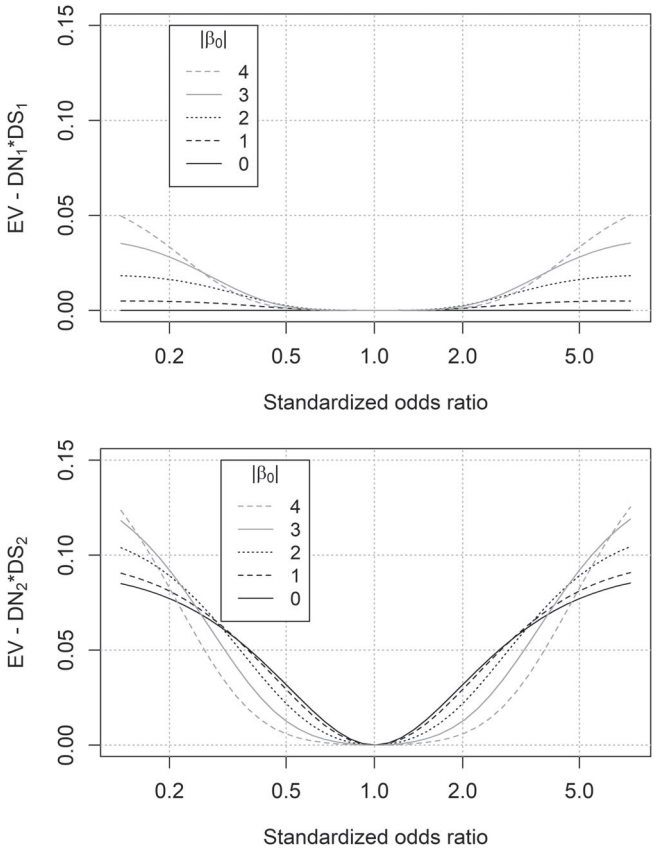

FIGURE 3 Discrepancies between EV and $DN_1 \cdot DS_1$ (top panel) and $DN_2 \cdot DS_2$ (bottom panel) for the logistic regression model (37) and a single standard normally distributed prognostic factor

Further empirical investigations have shown that the marginal distribution of X can also affect the magnitude of Δ_1 and Δ_2, for example, within the range of OR considered in Figure 3 we observed values of Δ_1 between -0.004 and 0.158 for log-normally distributed X. However, we refrain from presenting detailed results, and note that the linear predictor from a multivariable regression model often also follows a normal distribution.

Generally, for the typical effect sizes with most clinical and epidemiologic studies (OR between 0.2 and 5) Δ_1 is close enough to zero. In addition, also more extreme effects (eg, $OR > 50$ in the data set of the next section) tend to result in only negligible values for Δ_1. Therefore, the conceptually interesting view of explained variation as a product of the degrees of necessity and of sufficiency can be kept for variant 1 also for continuous prognostic factors.

6 | FULL ANALYSIS OF THE AUSTRIAN COVID-19 DATA

The Austrian Ministry of Health made available their internal "Datenplattform COVID-19" for research purposes related to covid-19 in Austria.[2] We extracted information on age group (<20, 20-34, 35-49, 50-64, 65-79, ≥80), sex, place of residence (on the level of 32 groups of districts), time of positive covid-19 test (calendar week), and date of covid-19-related death (if applicable) for 19 747 infected individuals. The data cover information from February 24th to July 22nd, 2020. The first three age groups were merged due to extremely low numbers of deaths. Places of residence were summarized as urban agglomerations (at least 200 000 inhabitants, that is, Vienna, Graz, and Linz) vs others. For the time between positive test and death the Thursday of the respective calendar week was assumed as starting date. We restrict the data set to the 17 035 infected individuals with potential follow-up of at least 6 weeks.

Our interest focused on the effect of age group, sex, place of residence and month of infection by SARS-CoV-2 on mortality. While further prognostic factors such as comorbidities as well as more detail on the severity of the course of disease (hospitalization and intensive care) would have been of interest, this information was not available. The information available, however, permits an impression of the use of the measures of necessity and of sufficiency for dichotomous, ordinal and nominal outcomes developed in the previous and the current paper. The application of these measures does not replace standard statistical analysis tools but complements them. Therefore, here as well as in other applications,

TABLE 2 Covid-19 disease after infection by SARS-CoV-2 in Austria (as of July 22, 2020)

Variable	Austrian pop. in millions[a] (1)	# Tested positive (% of (1)) (2)	# Deaths by covid-19 (% of (2)) (3)	Marginal OR for death by covid-19 (4)	Partial OR for death by covid-19[b] (5)	# Late[c] deaths by covid-19 (% of (2)) (6)	Partial OR for ordinal outcome[d] (7)
Age group							
<50	5.25	8645 (0.16%)	7 (0.08%)	0.12	0.12	3 (0.03%)	0.12
50–64	1.96	4678 (0.24%)	31 (0.7%)	1.0	1.0	16 (0.3%)	1.0
65–79	1.22	2198 (0.18%)	238 (10.8%)	18.2	17.5	153 (7.0%)	17.5
≥80	0.47	1514 (0.32%)	418 (27.6%)	57.2	62.3	262 (17.3%)	60.8
Month of diagnosis							
February, March	—	8663	297 (3.4%)	1.00	1.00	195 (2.3%)	1.00
April	—	6922	370 (5.3%)	1.59	0.96	220 (3.2%)	0.97
May, June	—	1450	27 (1.9%)	0.54	0.58	19 (1.3%)	0.59
Sex							
Female	4.52	8651 (0.19%)	309 (3.6%)	1.00	1.00	192 (2.2%)	1.00
Male	4.38	8384 (0.19%)	385 (4.6%)	1.30	2.02	242 (2.9%)	1.98
Place of residence							
Other	6.49	12 807 (0.20%)	430 (3.4%)	1.00	1.00	268 (2.1%)	1.00
Urban agglomerations[e]	2.41	4228 (0.18%)	264 (6.2%)	1.92	1.84	166 (3.9%)	1.79
Total	8.90	17 035 (0.19%)	694 (4.1%)			434 (2.5%)	

Note: Patients with potential follow-up of at least 6 weeks.
[a] Population as of January 1st, 2020.[11]
[b] Partial effects are adjusted by all other factors of the table.
[c] "Late" deaths occur between week 2 and 6 after diagnosis of covid-19.
[d] Levels of 3-category ordinal outcome analyzed by proportional odds logistic regression: survived, late death, early death (within first week after diagnosis of covid-19).
[e] Urban agglomerations exceed 200 000 inhabitants (ie, Vienna, Graz, and Linz).

calculation of *DN* and *DS* values typically is embedded into standard analyses, often by appropriate types of regression models.

6.1 | Description and standard statistical analysis of the Austrian data

A statistical summary of covid-19 in Austria is given by Table 2, including some results from multivariable models (none of the pair-wise interactions were included due to statistical non-significance taking into account multiplicity; goodness-of-fit according to Hosmer-Lemeshow is not rejected). All calculations were done using SAS 9.4.

We learn from column 2 that, for the total population, fewer than two individuals per 1000 inhabitants had been infected until 6 weeks before July 22nd, 2020. In agreement with international data,[3,4,5] the risk of dying from covid-19 for infected individuals is negligible for ages below 50 - less than 1 per thousand in Austria. For the remaining age groups 50-64, 65-79, ≥80, this risk increases to roughly 1%, 10%, and 25%, respectively. This confirms the substantial effect of advanced age on covid-19 mortality. The effects of all other potential prognostic factors available to us, though statistically significant (all $P < 0.001$), are of much lower magnitude, in particular if looking at unadjusted and adjusted estimates of odds ratios (columns 4 and 5): the adjusted odds for death are twice as high for men than for women, and almost twice as high for covid-19 patients in urban agglomerations than in smaller and more rural settlements. The adjusted odds ratios decline with later months of infections, likely due to increasing medical experience with treatment of covid-19 rather than due to an increase in the extensive testing. The intensity of the latter had been relatively constant over the time

	Marginal			Partial		
Variable	***EV***	DN_1	DS_1	***EV***	DN_1	DS_1
Age group (4 cat.)	0.170	0.932	0.166	0.183	0.614	0.148
Month of diagnosis (3 cat.)	0.003	0.253	0.013	0.001	0.000	0.001
Sex	0.001	0.123	0.005	0.014	0.000	0.008
Place of residence (2 cat.)	0.004	0.176	0.023	0.010	0.000	0.006
Full model	0.193	0.933	0.179	—	—	—

Abbreviation: cat., categories.

TABLE 3 Explained variation (*EV*), degree of necessity (*DN*), and degree of sufficiency (*DS*) for Austrian covid-19 data for dichotomous outcome "covid-19 related death"

frame of interest.[12] The testing had exceeded 1 million tests by the end of July. Also repeated prevalence studies based on polymerase chain reaction antibody testing of SARS-Cov-2 in representative random samples in Austria indicate a very low amount of missed infections.[13] Thus substantial bias of any estimated effects on covid-19 mortality can be ruled out.

Spearman correlations between the four prognostic factors are close to 0, with one exception: urban agglomerations, with 27% of the population, contributed at the beginning of the pandemic, in February and March, only 16% of new covid-19 cases, but this proportion increased to 26% in April and to 69% in May and June. More important, the proportion of patients ≥80 years is approximately twice as high among infected women compared to men (11.2% vs 6.5%) and in April compared to the remaining months of diagnosis (13.1% vs 6.1% in February and March and 5.8% in May and June). This explains the most striking differences between adjusted (partial) and unadjusted (marginal) odds ratios in Table 2.

In general, if individual survival times are available one would prefer the specialized methods for such outcomes to methods for dichotomous or for ordinal outcomes. Nevertheless, by realistically assuming that infected individuals without a recorded death within 6 weeks survived and by limiting the time horizon of interest to 6 weeks (resulting in type 1 censoring at 6 weeks), we may correctly analyze the data also by methods for dichotomous and for ordinal outcomes, and compare these approaches. For the ordinal outcome we classify survival times into a few, medically well interpretable ordinal outcome categories: survived, died between 2 and 6 weeks (see column 6 of Table 2), died within 1 week after diagnosis. Slightly less than half of the deaths observed occurred in the first week after positive testing.

A proportional odds logistic regression model seems plausible, since separate binary logistic models for death (early or late) vs survived and for early death vs late death or survived result in reasonably close estimates of the slope parameter (2.16 and 1.94, respectively). The last column of Table 2 shows that results by a proportional odds model agree very much with those from binary logistic regression for the dichotomous outcome.

6.2 | Analysis of the prognostic factors in terms of EV, DN, and DS

While in Table 2 the odds ratios from logistic regression for the four prognostic factors considered are reported, in Table 3 we present corresponding values of explained variation of the outcome (*EV*) and of the suggested measures of the degrees of necessity (DN_1) and of sufficiency (DS_1). Note that for each of the prognostic factors in Table 3 a single value for each measure permits intuitive and easy comparisons. By contrast, in Table 2 each non-reference category receives a separate odds ratio estimate which hampers comparability between the prognostic factors.

We learn from Table 3 that age group definitely is the only relevant factor, and that the modest marginal DN_1 values of the other factors vanish if adjusted by age group. This is consistent with the fact that the values of the full model do not substantially improve beyond the effect of age group. Furthermore, and consistent with the similarity of results of binary logistic regression and proportional odds logistic regression (columns 5 and 7 of Table 2) the *EV*, *DN*, and *DS* values for the latter model ($EV^{ord} = 0.148$, ${DN_1}^{ord} = 0.937$, and ${DS_1}^{ord} = 0.137$ for the full model) roughly agree with those reported by Table 3 for the dichotomous outcome. Finally, the multiplicative decomposition of *EV* into *DN* and *DS* (see Section 5.2) clearly indicates that it is the low degree of sufficiency (DS_1) of available factors which leads to low *EV* and thus makes individual predictions of death difficult.

For the categorical age variable it is easy to identify the protective and harmful range, introduced in Section 2, for the dichotomous outcome: The death rates of 0.08% and 0.7% in the two lower age groups (<65 years) are well below the overall rate of 4.1%, marking these two groups as "protective"; in contrast the two higher age groups (≥65 years) are marked as "harmful" with rates of 10.8% and 27.6%.

For the ordinal outcome the homogenous effect across categories k assumed under proportional odds implies consistent protective ranges across k (and thus also consistent harmful ranges). This is due to the fact that a proportional odds model leads to a parallel shift of conditional probability estimates across categories k. If, in addition, the estimated intercepts are ordered according to the order of categories, then the unconditional estimates of $P(D_{k+})$ are monotone with k. In practice, these two effects nearly compensate each other such that the cutpoint between conditional and unconditional estimates, and therefore between the protective and the harmful range, is nearly the same across k. Applying the proportional odds model to our data, both the cutpoint for $k = 1$, that is, death (early or late) vs survived, and the one for $k = 2$, that is, early death vs late death or survived, is at 65 years. Simulations indicate that the observed proximity of cutpoints across k is typical also for realistic settings with continuous prognostic factors (results not shown).

We now demonstrate how the relations presented in Section 5 are realized in our data. If the dichotomized age variable (<65 vs ≥65) is used as the only prognostic factor for death from covid-19, then $\mathrm{Var}_<(P(D|X)) = \mathrm{Var}_>(P(D|X)) = 0$ with $DN_1 = 0.930$, $DS_1 = 0.142$. By replacing the dichotomized factor by the age group variable containing four categories (see Table 2) the protective and the harmful range are unchanged as are $P(D) = 0.041$, $P_<(D) = 0.003$, $P_>(D) = 0.177$ and $\alpha = 0.218$. However, $\mathrm{Var}_>$ now amounts to 0.0068 if we distinguish patients aged 65 to 79 (with death rate 10.8%) from those at least 80 years of age (with death rate 27.6%). Thus, DS_1 increases to 0.166 by using Equation (26). The increase in DN_1 is much smaller since the variance $\mathrm{Var}_<$ induced by dividing the protective range at age 50 is negligible (death rates of 0.08% vs 0.7%). Replacing the dichotomous age variable by the 4-category variant increases the discrepancy between EV and the product of DN_1 and DS_1 from 0 to $\Delta_1 = 0.015$. By contrast, DN_2 and DS_2 do not change, since they are independent of $\mathrm{Var}_<$ and $\mathrm{Var}_>$ (see Equations (27) and (28)). Accordingly, $\Delta_2 = 0.092$ is substantially larger.

We have so far considered the outcome as either dichotomous (survived, died) or as ordinal (survived, died within 2-6 weeks after positive test, died within the first week). If we are not convinced that prolonged survival for a few weeks and mostly under intensive care is to be preferred to a quick death within a week, we could consider the polytomous outcome as nominal, that is, without an assumed order. Such situations lead to gray zones between the preference for ordinal or for nominal scales.[14] For the covid-19 data the application of the nominal measures of Section 4 (with reference category "survived") results in very similar values ($EV^{nom0} = 0.139$, ${DN_1}^{nom0} = 0.933$, ${DS_1}^{nom0} = 0.112$ for the full model), due to numerical dominance of the reference category.

The estimates of explained variation for all outcomes considered remind us again that little is determined by age group on an individual level. All types of analysis show good numerical agreement, mainly because of the dominant proportion of survivors.

7 | DISCUSSION

We have argued that in many fields of study, especially empirical sciences, the concept of a cause being necessary or sufficient for an effect is nuanced. We have proposed measures DN and DS that quantify the degrees of necessity and sufficiency of a prognostic factor when the outcome is dichotomous, nominal, or ordinal. This allows, we hope, a more informative exploration of the relationship between "explanatory" and outcome variables. As we have seen, there is a close relationship between DN and DS and established measures of explained variation EV. Explained variation permits comparisons of the *importance* of individual prognostic factors,[15] also if measured on different scales and of different types (dichotomous, continuous, or qualitative), and even the comparison of groups of related factors.[16] By supplementing EV with DN and DS we can determine for example whether low EV is a result of low DN, low DS, or low DN and DS.

Tables 2 and 3 were obtained as part of our analyses of the Austrian covid-19 data. We considered four potential prognostic factors and confirmed the dominant role of age, with some marginal effects of the others, but no real partial effects after age was allowed for.

Note that among the four potential prognostic factors available in our Austrian covid-19 data only the place of residence qualifies as a modifiable risk factor. However, even here a causal interpretation of DN in the sense of an attributable risk (and DS as reverse attributable risk[1]) as well as a literal interpretation of the "harmful" (urban agglomerations) and "protective" (other place of residence) range necessitate a careful multivariable analysis also considering further potential prognostic factors such as comorbidities. Nevertheless, whether prognostic factors are seen as modifiable or not we regard DN and DS as a useful additional characterization of them. With our covid-19 data they tell us that total explained variation and thus predictability of individual outcomes can only be increased markedly by the addition of highly sufficient factors, presumably certain comorbidities, to our existing multivariable model.

In order to simplify application R functions and SAS macros for calculating *DN* and *DS* are provided at https://cemsiis.meduniwien.ac.at/en/kb/science-research/software/statistical-software/.

ACKNOWLEDGEMENTS

We gratefully acknowledge permission to get access to the Austrian Datenplattform COVID-19 by Gesundheit Österreich, and helpful discussions on Austrian covid-19 data with Michael Hummer, PhD (Gesundheit Österreich).

DATA AVAILABILITY STATEMENT

The data that support the findings of this study are available upon request from Gesundheit Österreich.[2] Restrictions apply to the availability of these data, which were used under license for this study.

ORCID

Andreas Gleiss https://orcid.org/0000-0001-8010-7665

REFERENCES

1. Gleiss A, Schemper M. Quantifying degrees of necessity and of sufficiency in cause-effect relationships with dichotomous and survival outcomes. *Stat Med*. 2019;38(23):4733-4748.
2. Austrian Federal Ministry of Social Affairs, Health, Care and Consumer Protection. Unpublished data obtained from the "Datenplattform COVID-19 Gesundheit Österreich". Accessed July 27, 2020.
3. Centers for Disease Control and Prevention. Demographic Trends of COVID-19 cases and deaths in the US reported to CDC. https://covid.cdc.gov/covid-data-tracker/#demographics. Accessed September 3, 2020.
4. Office for National Statistics. Deaths involving COVID-19, England and wales: deaths occurring in June 2020 - Characteristics of those dying from COVID-19. https://www.ons.gov.uk/peoplepopulationandcommunity/birthsdeathsandmarriages/deaths/bulletins/deathsinvolvingcovid19englandandwales/deathsoccurringinjune2020#characteristics-of-those-dying-from-covid-19. Accessed September 3, 2020.
5. Robert Koch Institute. Deaths due to Covid-19 by age and sex in Germany (in German). cited via statista. https://de.statista.com/statistik/daten/studie/1104173/umfrage/todesfaelle-aufgrund-des-coronavirus-in-deutschland-nach-geschlecht/. Accessed September 3, 2020.
6. Schemper M. Predictive accuracy and explained variation. *Stat Med*. 2003;22(14):2299-2308.
7. Schemper M, Henderson R. Predictive accuracy and explained variation in cox regression. *Biometrics*. 2000;56(1):249-255.
8. Woodward M. *Epidemiology: Study Design and Data Analysis*. Boca Raton, FL: CRC Press; 2014.
9. Oken M, Creech R, Tormey D, et al. Toxicity and response criteria of the eastern cooperative oncology group. *Am J Clin Oncol*. 1982;5:649-655.
10. Agresti A. *Analysis of Ordinal Categorical Data*. New York, NY: John Wiley & Sons; 1984.
11. Statistics Austria. Austrian population by age and sex as of Jan 1st, 2020 (in German). https://www.statistik.at/web_de/statistiken/menschen_und_gesellschaft/bevoelkerung/bevoelkerungsstruktur/bevoelkerung_nach_alter_geschlecht/index.html. Accessed September 3, 2020.
12. Austrian Broadcasting Corporation (ORF). Corona virus in Austria (data and maps; in German). https://orf.at/corona/stories/daten/. Accessed September 3, 2020.
13. Statistics Austria. SARS-CoV-2 prevalence study (in German). https://www.statistik.at/web_de/statistiken/menschen_und_gesellschaft/gesundheit/covid19/123051.html. Accessed September 3, 2020.
14. Van Calster B, Van Belle V, Vergouwe Y, Timmerman D, Van Huffel S, Steyerberg EW. Extending the c-statistic to nominal polytomous outcomes: the polytomous discrimination index. *Stat Med*. 2012;31(23):2610-2626.
15. Healy MJR. Measuring importance. *Stat Med*. 1990;9:633-637.
16. Gleiss A, Zeillinger R, Braicu EI, Trillsch F, Vergote I, Schemper M. Statistical controversies in clinical research: the importance of importance. *Ann Onc*. 2016;27:1185-1189.

How to cite this article: Gleiss A, Henderson R, Schemper M. Degrees of necessity and of sufficiency: Further results and extensions, with an application to covid-19 mortality in Austria. *Statistics in Medicine*. 2021;40:3352–3366. https://doi.org/10.1002/sim.8961

APPENDIX A.

Here we present in detail how the calculation of the measures *DN* and *DS* for survival outcomes specializes to the calculation of DN^{ord} and DS^{ord} for ordinal outcomes, given in Section 3.

We start with survival outcomes and precisely follow notation and definitions in our previous paper (Section 3)[1]: in a given sample, let t_i, η_i, and x_i denote observation time, censoring indicator, and vector of prognostic factors, respectively, for individual i ($1 \le i \le n$). Assume there are m distinct survival times in the sample, at times $t_{(j)}$ ($1 \le j \le m$), with d_j deaths at $t_{(j)}$. Then at each distinct death time $t_{(j)}$ we estimate the degree of necessity as

$$\widehat{DN}_{1,t_{(j)}} = \left[\frac{1}{n_<(t_{(j)})} \sum_{i=1}^{n} I(\hat{F}(t_{(j)}\,|\,x_i) < \hat{F}(t_{(j)})) \left(\frac{\hat{F}(t_{(j)}) - \hat{F}(t_{(j)}\,|\,x_i)}{\hat{F}(t_{(j)})} \right)^2 \right]^{1/2} \tag{A1}$$

with $I(.)$ denoting the indicator function, $\hat{F}(t)$ and $\hat{F}(t|X)$ unconditional and conditional cumulative distribution functions, and $n_<(t_{(j)})$ the number of subjects i with $\hat{F}(t_{(j)}|x_i) < \hat{F}(t_{(j)})$ at time $t_{(j)}$.

An overall estimate of DN_1 is obtained via a weighted average of the $\widehat{DN}_{1,t_{(j)}}$ over survival times, with weights designed to compensate the attenuation in observed death due to earlier censorship:

$$\widehat{DN}_1^{surv} = w^{-1} \sum_{j=1}^{m} \hat{G}(t_{(j)})^{-1} d_j \widehat{DN}_{1,t_{(j)}} \tag{A2}$$

with $w = \sum_{j=1}^{m} \hat{G}(t_{(j)})^{-1} d_j$, and $\hat{G}(t)$ denoting the Kaplan–Meier estimator of the censoring or "potential follow-up" distribution.

Moving to ordinal categorical outcomes interpreted as uncensored survival outcomes, first, the function $\hat{G}(t)$ always assumes values of 1 and therefore can be dropped. Second, to comply with notation of Section 3, we replace the m distinct times $t_{(j)}$ ($1 \le j \le m$) with d_j deaths at $t_{(j)}$ by $K+1$ ordinal categories ($0 \le k \le K$), $K = m-1$, with estimated probabilities $\hat{p}_k = d_{k+1}/w$, $\sum_{k=0}^{K} \hat{p}_k = 1$.

Then we obtain

$$\widehat{DN}_1^{surv} = \sum_{k=0}^{K} \hat{p}_k \left[\frac{1}{n_<(k+)} \sum_{i=1}^{n} I(\hat{F}(k\,|\,x_i) < \hat{F}(k)) \left(\frac{\hat{F}(k) - \hat{F}(k\,|\,x_i)}{\hat{F}(k)} \right)^2 \right]^{1/2} \tag{A3}$$

with $n_<(k+)$ denoting the number of subjects i with $\hat{F}(k|x_i) < \hat{F}(k)$ in category k.

Now, in the presentation of the measures for survival outcomes,[1] large survival times are considered favorable, while with ordinal outcomes large values are considered unfavorable (without loss of generalizability). To make both approaches comparable we transform survival times T to ordinal categories Y in such a way that large (small) survival times receive favorable (unfavorable) outcome scores. Therefore, the longest survival times are scored best, that is, are moved to category "0," the longest but one survival times are moved to the best but one category "1," … , and finally, the shortest survival times are moved to the "worst" category "K." As a consequence $\hat{F}(k) = \sum_{j=0}^{k} \hat{p}_j$ in (A3) turns into $\hat{S}(K-k) = \sum_{j=K-k}^{K} \hat{p}_{K-j} = \sum_{j=0}^{k} \hat{p}_j$ in (A4) below; analogously for $\hat{F}(k|x_i)$ and $\hat{S}(K-k|x_i)$.

Furthermore, note that under no censoring $\hat{F}(K) = 1$, $\hat{F}(K|x_i) = 1$ resulting in $\hat{F}(K) - \hat{F}(K|x_i) = 0$ in (A3) which turns into $\hat{S}(0) - \hat{S}(0|x_i) = 0$ in (A4). Therefore, in the version for ordinal outcomes (A4) averages are taken over all categories except 0 (ie, over K components), which always would contribute a value of 0. With survival outcomes averages (A2) are taken over all distinct survival times (ie, over $m = K+1$ components), which is necessary for the case of censoring. Without censoring the component related to the longest survival time contributes 0 to the averaging and therefore, correctly, should not be included. With the typically observed many distinct survival times the downward bias from including it in the estimation of DN and DS should be small. With ordinal outcomes, however, often having as few as 3 to 5 categories, it is essential not to include category 0 in the averaging. This results in

$$\begin{aligned} \widehat{DN}_1^{ord} &= \left(\sum_{k=0}^{K-1} \hat{p}_{K-k} \right)^{-1} \\ &\quad \cdot \sum_{k=0}^{K-1} \hat{p}_{K-k} \left[\frac{1}{n_<((K-k)+)} \sum_{i=1}^{n} I(\hat{S}(K-k\,|\,x_i) < \hat{S}(K-k)) \left(\frac{\hat{S}(K-k) - \hat{S}(K-k\,|\,x_i)}{\hat{S}(K-k)} \right)^2 \right]^{1/2} \\ &= \left(\sum_{k=1}^{K} \hat{p}_k \right)^{-1} \sum_{k=1}^{K} \hat{p}_k \left[\frac{1}{n_<(k+)} \sum_{i=1}^{n} I(\hat{S}(k\,|\,x_i) < \hat{S}(k)) \left(\frac{\hat{S}(k) - \hat{S}(k\,|\,x_i)}{\hat{S}(k)} \right)^2 \right]^{1/2} \end{aligned}$$

which is (9) of Section 3. The relationships for DS^{ord} and EV^{ord} can be presented along the same lines.

APPENDIX B.

For proving (24) we start from its right-hand side which is equal to

$$\begin{aligned}
&(1-\alpha)\mathrm{E}_<(\mathrm{P}(D\,|\,X)-\mathrm{P}_<(D))^2+\alpha\mathrm{E}_>(\mathrm{P}(D\,|\,X)-\mathrm{P}_>(D))^2+\alpha(1-\alpha)[\mathrm{P}_>(D)-\mathrm{P}_<(D)]^2\\
&\quad=(1-\alpha)[\mathrm{E}_<(\mathrm{P}(D\,|\,X))^2-\mathrm{P}_<(D)^2]+\alpha[\mathrm{E}_>(\mathrm{P}(D\,|\,X))^2-\mathrm{P}_>(D)^2]+\alpha(1-\alpha)[\mathrm{P}_>(D)^2-2\mathrm{P}_>(D)\mathrm{P}_<(D)+\mathrm{P}_<(D)^2]\\
&\quad=[(1-\alpha)\mathrm{E}_<(\mathrm{P}(D\,|\,X))^2+\alpha\mathrm{E}_>(\mathrm{P}(D\,|\,X))^2]-[\alpha^2\mathrm{P}_>(D)^2+2\alpha(1-\alpha)\mathrm{P}_>(D)\mathrm{P}_<(D)+(1-\alpha)^2\mathrm{P}_<(D)^2]\\
&\quad=\mathrm{E}(\mathrm{P}(D\,|\,X))^2-[\alpha\mathrm{P}_>(D)+(1-\alpha)\mathrm{P}_<(D)]^2\\
&\quad=\mathrm{E}(\mathrm{P}(D\,|\,X))^2-\mathrm{P}(D)^2\\
&\quad=\mathrm{Var}(\mathrm{P}(D\,|\,X))
\end{aligned}$$

APPENDIX C.

Some properties of $\Delta_1 = EV - DN_1 \cdot DS_1$:

- If $\mathrm{Var}_< = \mathrm{Var}_>$ then $\Delta_1 \le 0$ since

$$EV^2 - DN_1^2 \cdot DS_1^2 = -(2\alpha-1)^2[\mathrm{P}_>(D)-\mathrm{P}_<(D)]^2\mathrm{Var}_</[\mathrm{P}(D)(1-\mathrm{P}(D))]^2 \le 0$$

- If $\mathrm{Var}_< = \mathrm{Var}_>$ and $\alpha = 0.5$ then $\Delta_1 = 0$ following from (33), (34), and (35).
- A slightly more general, but still only sufficient condition for $\Delta_1 = 0$ is given by

$$\mathrm{Var}_>(\mathrm{P}(D\,|\,X)) = \mathrm{Var}_<(\mathrm{P}(D\,|\,X)) + (2\alpha-1)[\mathrm{P}_>(D)-\mathrm{P}_<(D)]^2$$

  This results from setting the expressions under the square roots of (33) and (34) to be equal, in which case their product equals the numerator of (35). However, with a logistic regression model and normally distributed X, this condition is only fulfilled for the case $\mathrm{Var}_< = \mathrm{Var}_>$ and $\alpha = 0.5$ (scenario $\beta_0 = 0$; see the empirical investigation in Section 5.2).
- If $\mathrm{Var}_< = \mathrm{Var}_> = 0$, that is, X is dichotomous, then $\Delta_1 = 0$ independent of α.