Letters to the Editor

Predicting complications in interventional cardiology

The article by Budde et al.[1] gives rise to the hope that coronary angioplasty complications can be predicted, almost with certainty, in individual patients. In this paper, the authors classify their patients with an accuracy of 95%, into groups with low (<10%), moderate (70–90%) and high (>90%) probabilities of complications.

The rules of classification depicted in Fig. 1 in this paper derive detailed analysis. One of the first steps in the decision-tree is the identification of NYHA <III patients with directional atherectomy. They are divided in two groups: those with and without diuretics at low and high risk of complications. Table 2 in the paper tells us that 12 patients were treated with atherectomy. Following the authors’ definition of low and high risk as <10% and >90% of complications, their claim to classify the patients with >95% accuracy would mean that virtually every patient receiving diuretics had complications from atherectomy and every patient without diuretic had no such problems. Do diuretics in atherectomy patients really matter that much? One could look at this from another point of view: the authors extracted the classification rules from 2500 items per patient entered on a computer. If one looks at the distribution of such a mass of data in groups of patients with and without complications, imbalances with respect to several of these are to be expected (statistically significant) differences will be observed in 25 of these items with formal P values <0.01). The association of diuretics with atherectomy complications is, therefore, much more likely to be a matter of chance than reflecting a clinically relevant relationship. One would not expect subsequent patients with and without diuretics to be similarly different in their complication rates after an atherectomy procedure. Should we then trust this rule to predict complications in an individual patient?

This problem is even more evident with the next step of the decision tree, grouping patients with rotablator therapy into (a) eccentric stenosis (=high risk), (b) concentric stenosis <1 cm (=low risk) and, (c) concentric stenosis >1 cm (high risk). Only five patients were treated with a rotablator (Table 2). That means that the largest of these three groups consisted of three patients at most, and the smallest was only a single patient. This is close to the popular ‘statistical’ assumption that (with respect to the occurrence of an event) once was ‘never’ and twice was ‘ever’. How would these rules be changed by artificial intelligence if the next two patients with a concentric stenosis >1 cm had no complications after rotablator treatment (which is a realistic assumption)? We should not be too confident about the predictions of systems such as INTERVENT which are used to guide therapeutic strategy decision making in individual patients.

Let us look at a more relevant group, classified as high-risk by the computer algorithm: female patients >75 years of age being treated for unstable angina. The ALKK registry[2] presently contains 3107 patients with these characteristics, of whom 309 had complications (death, MI, vessel closure, repeat PTCA or bypass surgery). The aPTT values of these patients are unknown to us but many will have been on iv heparin in their unstable status. The 10% complication rate in these patients is clearly higher than average, but far from the definitive prediction of complications in each case, as claimed by the authors.

On the other hand, if the predictions were reliable, this would mean that complications resulting from an intervention could be foreseen almost with certainty? The only plausible consequence then would be to abstain from that specific intervention, except in some rare and desperate situations. Fortunately, in the vast majority of patients scheduled for interventional procedures, the probability of having no complications is far greater than 50%, but no group can be defined as being at high risk in the definition of the INTERVENT investigators. As pointed out in the editorial[2] accompanying the article, we should probably prefer to visit the caves of Cumae before relying on artificial intelligence overruling plausibility and experience in everyday clinical practice.

References


A reply

‘We achieve more than we know. We know more than we understand. We understand more than we can explain’

(Claude Bernard)

The letter by Vogt indicates some misunderstanding of the basic techniques presented in our paper, in particular the fact that in his three ‘counter-examples’ only one rule is used for a classification. It is well known that neither by single nor multiple parameters (like the ABC criteria for classification of stenoses) nor single rules (If . . . then . . . else) can an individual outcome be calculated exactly. Therefore by looking at one or more parameters or one rule (e.g. if sex=female, unstable angina=yes, ostial lesion=no, aPTT≥40s=high-risk for post-lab complications with probability 0.6) it is easy to find counter-examples.

The knowledge discovery process for calculating the individual outcome of patients presented in our article consists of various technical steps to construct (called boosting technique [Freund, 1997[1]] a set of classifiers (Ripley 1996[2]) of which we presented just one as an example. In practice, when a case is classified, all these classifiers are consulted before a decision is made. On an average, 10 to 15 different rules have to be evaluated