

Supplementary Material

Optimal Biomarkers Design for Drug Safety Evaluation Using Microelectrode Array Measurements

1 Imperfect electrode model

The bidomain model describes the evolution of the transmembrane potential V_m and the extracellular potential ϕ_e in a domain Ω . We denote by R_i , R_{el} and C_{el} , the internal resistance of the measurement device, the electrode resistance and the electrode capacitance respectively. The field potential ϕ_f^k measured on an electrode e_k is given by $\phi_f^k = R_i I_{el}^k$, where I_{el}^k is linked to the averaged extracellular potential $\phi_{e,mean}^k$ at the electrode e_k by the equation:

$$\frac{dI_{el}^k}{dt} + \frac{I_{el}^k}{\tau} = \frac{C_{el}}{\tau} \frac{d\phi_{e,mean}^k}{dt}, \quad (1)$$

where $\tau = (R_i + R_{el})C_{el}$.

For the present study the parameters values are summarized in Table 1.

C_{el}	R_i	R_{el}
$1\mu F$	$2M\Omega$	$10M\Omega$

Table 1: Parameters used for the imperfect electrode model.

2 Biomarker computation

In this section, we provide details on how to compute the biomarkers from FP time series. For a given signal, we denote by t the time vector and y the FP. Each signal is divided in two parts: the depolarization region (t_1, y_1) and the repolarization region (t_2, y_2) as shown in Figure 1.

Depolarization amplitude (DA) The DA is simply defined as the difference between the maximum and minimum values of the potential during the depolarization:

$$DA = \max(y_1) - \min(y_1). \quad (2)$$

Repolarization amplitude (RA) The RA is defined as the maximum (in absolute value) of the repolarization.

$$RA = \max(|y_2|). \quad (3)$$

Field potential duration (FPD) The FPD is defined as the time difference between the maximum (in absolute value) of the depolarization and the maximum (in absolute value) of the repolarization. Let $t_d = t \left[\operatorname{argmax}_t (|y_2(t)|) \right]$ and $t_r = t \left[\operatorname{argmax}_t (|y_1(t)|) \right]$. Then,

$$\text{FPD} = t_r - t_d. \quad (4)$$

Area under the curve of the repolarization wave (AUC) The AUC is defined as the area under the curve of y_2 truncated around $\pm\Delta t$ of t_r . We used $\Delta t = 100\text{ms}$. The integral is approximated using the trapezoidal rule.

$$\text{AUC} = \left| \int_{t_r - \Delta t}^{t_r + \Delta t} y_2(t) dt \right| \quad (5)$$

Repolarization center (RC) The RC is defined as the offset of the barycenter (with respect to time) of the repolarization wave.

$$\text{RC} = \frac{1}{\text{AUC}} \int_{t_r - \Delta t}^{t_r + \Delta t} t y_2(t) dt - t_r \quad (6)$$

Repolarization width (RW) The RW is defined as the normalized standard deviation of the repolarization wave.

$$\text{RW} = \frac{1}{\text{AUC}} \left[\int_{t_r - \Delta t}^{t_r + \Delta t} t^2 y_2(t) dt - \left(\int_{t_r - \Delta t}^{t_r + \Delta t} t y_2(t) dt \right)^2 \right]^{1/2}. \quad (7)$$

Field potential notch (FPN) The FPN is defined as the potential value 4ms after t_d . The FPN value is smoothed out by multiplying the signal with a test function and then integrat the product. This proves to be less sensitive to noise than just a point-wise evaluation. Let $\phi(t_1) = \exp \left[-\frac{(t_1 - (t_d + 4))^2}{.04} \right]$. Then,

$$\text{FPN} = \int_{t_1} y_1(t_1) \phi(t_1) dt_1. \quad (8)$$

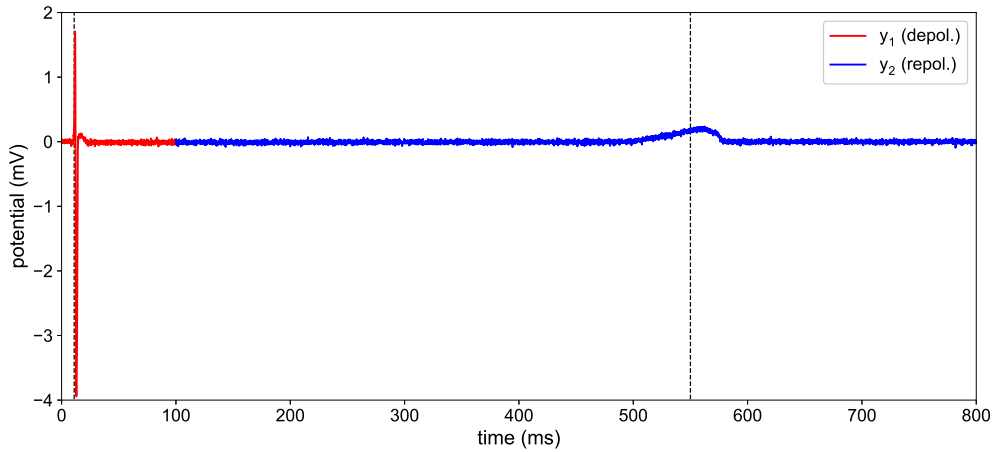


Figure 1: FP depolarization and repolarization regions.

3 Classification metrics

We now present the two different classification metrics used in this work.

Cohen’s kappa Cohen’s kappa, denoted by κ , is particularly suited for multi-class and/or imbalanced classification problems. The main idea is that it measures the labeling discrepancy between two annotators (or classifiers). It is simply adapted to our case by considering one of the annotators as the ground truth (true labels). Its formula reads:

$$\kappa = \frac{p_o - p_e}{1 - p_e}, \quad (9)$$

where p_o is the observed agreement between the two annotators and p_e is the probability of an agreement between two random annotators. For further details, the reader is referred to Scikit-learn’s implementation ¹ of Cohen’s kappa.

ROC The receiver operating characteristic area under curve (ROC AUC, later referred to as AUC for the sake of clarity) is basically associated with binary classification problems. In our case, one can define a AUC for each class k by considering all the other classes as only one class. With SVC it is possible to evaluate, in addition to the predicted class, the probability of belonging to each class. Given a threshold parameter (that varies between 0 and 1), it is possible to decide if a sample belongs to a given class when the SVC probability returned for this class is greater than the threshold parameter. The predicted class therefore depends on this parameter. When all samples of the validation set have been tested, the following quantities are computed, for each class k and for a given threshold parameter:

- true positives (TP): number of samples affected to class k which are actually in class k .
- false positives (FP): number of samples affected to class k which are actually *not* in class k .
- true negatives (TN): number of samples affected to another class than k which are actually in class k .
- false negatives (FN): number of samples affected to another class than k which are actually *not* in class k .
- true positive rate (TPR): $\frac{TP}{TP+FN}$.
- false positive rate (FPR): $\frac{FP}{FP+TN}$.

The ROC is the curve of TPR against FPR as the threshold parameter varies between 0 and 1. The AUC is simply the area under this curve.

¹<https://github.com/scikit-learn/scikit-learn/blob/ab93d65/sklearn/metrics/classification.py#L278>