

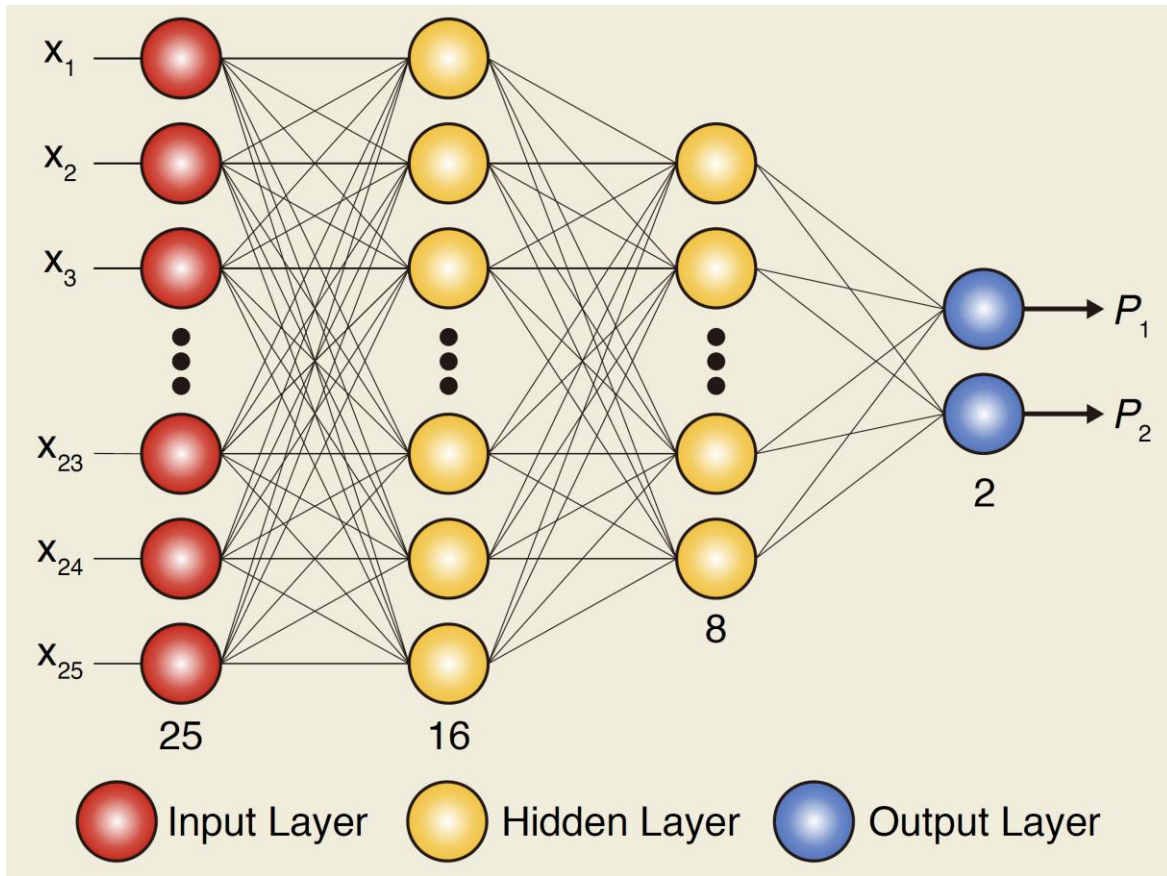


Fig. S6. Developing the DrABC Model through the Hierarchical Neural Network.

The prediction model was developed using the clinical characteristics in the discovery cohort and validated in the multi-center validation cohort. It starts with an input layer that contains 25 neurons (Table S2), which corresponds to features associated with the carriers of germline pathogenic variants in cancer predisposition genes. Then two hidden layers follow, with 16 and 8 neurons, respectively. Furthermore, a non-linear activation function, namely Scaled Exponential Linear Unit (SELU) (29), is attached to hidden layers' outputs, and it helps maintain the representation distributions to be close to unit Gaussian. Finally, we have an output layer with 2 neurons, which

comes with a sigmoid activation function, so that it produces two valid probabilities (i.e., in the range of $[0,1]$): P_1 and P_2 .