Abstract: Patients undergoing radiological diagnostic/screening procedures are often assigned multiple raters to assess their likelihood of disease based on the resultant medical images. For example, computed tomography (CT) and/or magnetic resonance (MR) are typically applied to patients visiting the emergency department with acute abdominal pain, and the images produced are interpreted by several radiologists for the presence of appendicitis. Proper combination of the ratings from different raters/modalities can increase the diagnostic accuracy compared with one from a single rater/modality alone. We develop a nonparametric statistical procedure to estimate/learn the optimal rule of combining multiple ratings based on a set of training data with known disease status. Under reasonable assumptions on the expertise of the raters and conditional independence between the raters given certain latent disease characteristics, we show that the risk score, i.e., the probability of disease given the ratings, is a monotonic function in each rating with the others held constant, whether the training data are collected from cohort or case-control designs. With this knowledge, we set out to estimate the nonparametric multivariate monotone risk score function using a likelihood-based EM-type algorithm. The estimated risk score function then serves as the basis for combined diagnosis. In simulation, the new method outperforms standard methods such as parametric modelling and the support vector machine for classification. As an illustration, we analyze real data collected on a group of potential appendicitis patients using the proposed methodology.