Purpose:
The Lung Image Database Consortium (LIDC) is an ongoing multi-institutional study funded by the NIH to collect and evaluate volumetric image data of tumors of the lung. A primary objective of this study is to ascertain a definitive nodule-detection ground-truth for future training and testing of competing computer-assisted detection algorithms. However, a great disparity exists among the (four) readers' interpretations preventing an objective determination of ground truth. The goal of the current work is to apply statistical methodologies to determine the maximum-likelihood nodule-detection ground truth for the LIDC data.

Methods:
Our method is motivated by the Simultaneous Truth and Performance Level Estimation approach by Warfield et al. and is a specific implementation of iterative, expectation-maximization and maximum likelihood (EM-ML) steps tailored for the nodule detection task. A key, necessary preliminary step is the ranking of readers, on a per patient basis, by the number of nodules detected; that is, their relative aggressiveness in detection.

Results:
Application of the EM-ML iterative scheme resulted in a reasonable estimate of reader performance and ground-truth that was obtained consistently for a variety of initial states. The converged ground truth conveniently matched identically the three out of four consensus from the LIDC readers, independent of stratification by aggressiveness.

Conclusions:
The application of statistical methods to the problem of nodule detection from multiple readers in the absence of a known ground-truth has been implemented on the LIDC outcomes to produce a reasonable and objective determination of ground-truth. This ground-truth permits the assessment of the performance of the LIDC readers based on relative aggressiveness and enables the LIDC detection data to be used for future CAD evaluation.