May 7, 2020

Submitted via Regulations.gov

NIOSH Docket Office
Robert A. Taft Laboratories, MS-C34
1090 Tusculum Avenue
Cincinnati, OH 45226

Re: Docket No. CDC-2019-0088 NIOSH-330– Coal Workers' Health Surveillance Program: B Reader Decertification and Autopsy Payment; Comments of the American College of Radiology

The American College of Radiology (ACR)—a professional association representing nearly 40,000 diagnostic radiologists, interventional radiologists, radiation oncologists, nuclear medicine physicians, and medical physicists —appreciates the opportunity to provide comments in response to the B Reader Decertification proposed rule. We concur with NIOSH's premise that there are clear examples of unreasonably inaccurate radiograph classifications in formal litigation and compensation proceedings which may be suggestive of B-Reader bias or skill-deficits. We therefore agree that there is a need for an enhanced quality assurance (QA) program – beyond the B-Reader exam process – to ensure the integrity of B-Reader interpretations. We have some concerns that the proposal detailed in the NPRM could invite abuse and result in inappropriate decertification of competent B-readers if implemented without appropriate safeguards. Moreover, the proposed rule does not address other important quality assurance needs that are critical to diagnostic accuracy and we respectfully urge NIOSH to incorporate these additional quality measures in this or future rulemakings. Our specific concerns and suggestions for addressing them are detailed below.

ILO standards do not necessarily translate to the appropriate clinical or medicolegal contexts
The *General Instructions in the Guidelines for the Use of the ILO International Classification of Radiographs of Pneumoconiosis Revised Edition 2011*, specifically state that, " In epidemiological studies, therefore, the study protocol will usually require that <u>all</u> appearances described in these guidelines and seen on standard radiographs are to be classified…When the Classification is used for some clinical purposes, the protocol may require that medical readers classify <u>only those appearances which the reader believes or suspects to be pneumoconiotic in origin</u>." (underline added for emphasis). It is notable that other lung diseases may mimic pneumoconiosis. Thus, the application of the ILO classification in a blinded fashion is appropriate for screening (as is performed in the CWHSP) but in specific clinical or medicolegal instances, additional medical or exposure information may inform readings and thus trigger the second instance – for example, a miner with sarcoidosis, connective tissue disease or drug toxicity that should have a work-up and receive therapy for these treatable conditions. The final rule, and NIOSH's implementation of the rule, should account for this divergence of ILO applications in screening, clinical, and medicolegal standards.

NIOSH should consider using thin section non-contrast chest CT as gold standard, rather than a B-reader panel to assess chest radiographic interpretation

NIOSH's decertification proposal would apply to a "reader found by NIOSH to have engaged in a pattern of providing unreasonably inaccurate chest radiograph classifications in practice those that are found by the Program to diverge substantially from a competent interpretation of the radiographs, as determined by a panel of practicing certified B Readers selected by NIOSH." However, the proposed rule does not indicate the criteria NIOSH will use to select this panel. Without radiologists on staff at NIOSH, it is not clear how NIOSH will assess the qualifications of prospective panel members relative to the B-Readers who are being assessed; prior published data indicates that radiologists outperform other specialties on the B-reader exam. The evaluation of interstitial lung disease is a complicated and ever evolving field often requiring subspecialists in the fields of pulmonary and radiology. The selection criteria NIOSH uses to identify the B-readers used to review these cases must be transparent and account for the complexity of ILD; the involvement of radiologists who have the opportunity to interpret a large volume of radiography and CT for ILD in the setting of multidisciplinary evaluation would be preferable. As in other ILDs, NIOSH should consider using thin section non-contrast chest CT as gold standard, rather than panel chest radiographic interpretation. This better follows the workup for all other ILDs. It would serve two purposes: to provide a definitive answer for the individual complaint and to provide a compendium of cases for future B-reader training.

The decertification proposal outlined in the NPRM raises a number of operational concerns:

a. **Potential for abuse** – The provision allowing "any interested party" to make a complaint could invite abuse by outside parties seeking to suppress expert witnesses. The lack of a process to exclude "interested" parties from submitting multiple unsubstantiated complaints could create system backlogs and undermine the reputational standing of competent B-Readers. As stated in the NPRM, the proposed process is intended to identify "systematic" abuse not a single episode of discrepancy.

b. **Likelihood of sampling bias** - Identifying examples from the court system introduces substantial sampling bias. The vast majority of cases reach settlement and a low percentage of cases end up in trial. Those that end up in court are, by definition, problematic interpretations, so even experienced interpreters may not agree. Perhaps a substantially higher bar to file a complaint should be established for cases identified in this manner or from attorneys.

c. **Proposal does not account for reader variability** - It is also notable that this process assumes willful misconduct rather than standard inter- and intra-reader variability; the chest radiograph is an interpretation NOT a lab test. Old studies of CXR interpretations estimate a 30% error rate with 10% missed findings and 20% over-diagnosis. This is for ALL CXR findings not ILD - one of the most challenging CXR diagnoses.

d. **The 3 case threshold is not appropriate for readers with high volumes** - This proposal does not take into account the volume of cases read by an experienced B-reader and the percentage of cases read correctly. To illustrate, suppose a B-reader interprets 5000 cases/per year and is correct 98% of the time (an inconceivably high expectation). The B-reader would be wrong not just 3 times per year but 100 times!

e. **The proposal should provide an opportunity for B-Readers to perform corrective action prior to decertification.** The NPRM provides: "To demonstrate a pattern of inaccurate classifications, the valid complaint must provide radiographs from three or more patients conducted within a one-year period that are determined by the Program to be inaccurate." This should be reworded as follows: "within a one year period following

the date of the finalized determination of the 1st valid complaint." The date of the 1st, 2nd and 3rd interpretation must be considered when decertifying a B-reader. That is, when a B-reader is issued the results of the 1st investigation and found to be inaccurate, the 2nd complaint should only be allowed to be filed on an interpretation issued after the date that the 1st investigation results were finalized in order to allow the B-reader to take corrective action. Similarly, the 3rd complaint should only be filed on an interpretation after the date that the 2nd investigation results are finalized.

f. **B-Readers should not be penalized for ambiguities in the ILO classification system.** Certain areas of the classification remain controversial. For instance, the ILO system has no instructions on the classification of pseudoplaque (nor does the pathology literature). Reasonable and competent B-readers may disagree on how these findings are categorized in the current classification.

Enhanced equipment and image quality standards are critical aspects of a B-Reader QA program

Finally, it is important to note that there are equal if not more overarching QA concerns that remain unaddressed in this proposed rule. The applicable Federal regulations establish minimum standards for digital radiography equipment; however, they fall far short of ensuring quality digital imaging for interpretation. Moreover, if an exam is not correctly performed, no amount of reader experience can ensure adequate interpretation. Digital imaging is particularly problematic because the impact of technique can be more difficult to recognize and is less well described than analog imaging. Low dose digital image acquisition may produce substantial image noise simulating small "p" opacities. Post-processing algorithms (especially edge-enhancing techniques) and some compression techniques can simulate pneumoconiosis. Equipment and radiographic technique certification is a necessary and fundamental component of this proposed rule that is missing. It is critical that NIOSH create a system to regulate the quality of digital radiography performed and establish a mechanism to identify facilities that are systematically creating low quality images. Certification by NIOSH physicists on control phantoms would be a great start.

As always, the American College of Radiology welcomes continued dialogue with NIOSH on matters of shared interest. Please contact Gloria Romanelli, JD, ACR Senior Director of Legislative and Regulatory Relations at (202) 223-1670/gromanelli@acr.org with questions or concerns.

Sincerely,

Geraldine B. McGinty, MD, MBA, FACR
Chair, Board of Chancellors
American College of Radiology