quality, reported and required exposure, and the image post-processing and appearance. Dr. Samei reviewed the physical bases of the technologies (sensitive layer, coupling layer, collection layer).

CR, developed in 1975, works by using x-rays to stimulate a layer containing phosphor, producing a latent image. Laser scanning then produces light emission, which is collected in the form of digital signals. The inherent image quality is governed by the lateral spread of laser light and the phosphor thickness. Flat panel technology was developed in the late 1990s and, although expensive, outperforms CR in quality and speed. It makes use of discrete pixel capacitors—and there are two types—photoconductor detection and phosphor detection. Charge-coupled systems, which also were developed in the 1990s, use phosphors in combination with discrete charge-coupled devices (CCDs) and metal-oxide semiconductor light sensors. Image quality is comparable to CR. Scatter, which reduces image quality, is an ever-present attribute of chest radiography, whether analog or digital. Dr. Samei explained the technique of slot-scanning DR, which can reduce scatter significantly.

Dr. Samei concluded that digital radiography offers advantages over film screen, that current technologies offer varying image quality, and that an initiative is needed to address the different systems and to unify conditions. He recommended that the radiography field:

- Standardize image acquisition and processing protocols
- Institute robust quality control and preventive maintenance programs
- Develop facility and equipment accreditation programs

CR and DR Chest Radiographic Image Parameters for the Pneumoconioses: The Japanese Approach and Experience – Narufumi Suganuma, M.D., Ph.D.

Dr. Narufumi Suganuma, of Kochi University Medical School, stated that because the Japanese Pneumoconiosis Law uses radiographic classification as a scale to determine administration class of dust-exposed workers and compensation, revision of the law is socially sensitive. A Japanese digital radiography taskforce began the revision process by defining appropriate digital

radiography parameters for classification. They built upon an earlier task force's recommended parameters for grayscale and spatial frequency and developed parameters for the Canon digital radiography system.

The task force conducted a reading trial of the classification of pneumoconiosis using a film-screen system compared to hard copy images from the same individual but produced using a digital radiography system. The study found that, for profusion, about 15 percent of cases were over-read by digital radiography and about 6 percent were under-read. About 80 percent were classified as the same.

Conclusions of this trial were that (1) there was crude agreement between digital radiography and FSR for pneumoconiosis classification, (2) inter-reader agreement for FSR and digital radiography were $\kappa = .6072$ and .6968 respectively, and (3) digital radiography can be considered to have a capability of classifying pneumoconiosis changes in the chest equal to that of FSR. The task force recommended a digital radiography grayscale imaging parameter equal to the previous parameter for FSR (1.6–2.0). It recommended that spatial frequency processing be turned off for digital radiography. The task force investigators subsequently studied other vendor systems (Philips, Siemens, GE, Toshiba, Hitachi, and Shimazu). They concluded that parameters for grayscale processing and spatial frequency processing must be standardized.

The task force concluded that digital radiography with appropriate settings could be used for legal management of patients with pneumoconiosis. The pre-storage parameter settings are more critical than the window level or width of stored image for the visualization of the appropriate image. It will not be practical to require that all CR or DR images be stored as raw data files. However, all digital radiograph data should be stored using a P-value as defined by DICOM. The use of soft copy (onscreen) images has yet to be evaluated, and users of soft copies are advised to consider appropriate variables for monitors and data.

Discussion

Dr. Yukinori Kusaka noted that readers in Japan tend to use hard copies and controls that were adopted over time. There is a need to establish standard parameters. Dr. Suganuma stated that a Japanese committee continues to evaluate parameters for different vendor monitor systems for quality assurance. Japan has adopted three control schemes and integrated them for calibration. Dr. Elizabeth Krupinski wondered why the Japanese study recommended no edge enhancement. Dr. Suganuma responded that this was because of the use of an analog standard (a preference).

Dr. Franzblau noted that a study being performed in Montana is collecting film and DR chest images, as well as CT images of pleural disease to be used as the gold standard. Dr. Fedotov wondered whether the diversity of systems will jeopardize the development of international standards. Dr. Samei stated that the solution is to have access to "for processing" data—that is, data that are ready for processing. It will also be important to know the system that produced the data, to have a gold standard, and to maintain a central Web server to load raw images. However, there are serious challenges to collecting raw image data.

Dr. David Clunie suggested recognizing and creating reference images that are appropriate to uses. He stated that it will be impossible to create an algorithm to consolidate data from various vendors/detectors. Yet, Dr. Samei noted we could readily adjust for such differences. Dr. Wagner noted that some hard-copy standards were developed using film stocks that have since been replaced. Dr. Petsonk proposed identifying parameters that are essential and creating minimum standards for them. Dr. Krupinski noted that the ACR guidelines provide such standards. They need to be adopted.

Dr. Daniel Henry cautioned that persons in far-flung places might have difficulty accessing a central site for standard references. Dr. Fred Prior suggested that we define a physics-based standard that vendors could apply to produce a certain quality. Dr. Eliot Siegel noted that experts in the field of digital mammography have proposed creating a harmonized raw data set and the use of a phantom that would be scanned to produce values for standards for data acquisition.

Image Presentation: Implications of Processing and Display – Michael Flynn, Ph.D.

Dr. Michael Flynn, of Henry Ford Health System, reviewed steps in processing images, which are used to transform digital radiographic data into display values for presentation at a workstation or film printer. A first processing step is preprocessing, in which raw data based on detected radiation energies are treated to create an image suitable for processing. The results are referred to as "for processing" data.

Display processing refers to subsequent steps in which "for processing" data are treated in five ways: grayscale rendition, exposure recognition, edge restoration, noise reduction, and contrast enhancement. Dr. Flynn described each of these processes. Grayscale rendition converts signal values to display values. Exposure recognition adjusts for high/low exposure. Edge restoration sharpens edges while limiting noise. Noise reduction features the reduction of noise while maintaining sharpness. Contrast enhancement entails increasing contrast to produce detail and produces the most dramatic and visible effects.

A final step, display presentation, refers to aspects related to the human visual system. Dr. Flynn reviewed the elements of viewing that affect the human interpretation of radiographic images. For example, the viewer is affected by viewing distance, display size, pixel size, and equivalent contrast, which refers to the role of brightness in the detection of contrast. Observer performance is best when the visual system is adapted to the average scene luminescence. Dr. Flynn listed the following display specifications: a luminance response of 350, a maximum brightness of 450 candelas per square meter or more, a pixel pitch of 0.210 mm or less, a diagonal size of 20–24 inches with a 4:3 or 5:4 aspect, and an ambient luminance that is less than 1 /4 of the minimum display luminance.

Dr. Flynn provided sample presentations of a chest image, showing, for example, the effects of tone-scale changes and edge restoration.

Discussion

In response to a question by Dr. Wagner, Dr. Flynn noted that, for edge restoration, the effects for nodules and irregular opacities are the same. The tradeoffs in adjusting parameters might be different for different detectors. Dr. Krupinski cited the factor of reader age in setting levels of enhancement for viewing. Dr. Lynch cited a need to study how processing affects the perception of pathology. The prettiest image might not be the most optimal in producing a perception of pathology. Perhaps we should develop digital standards with and without aggressive enhancement. Dr. Flynn envisioned a day when NIOSH and ILO support, based upon accepted observations, a standard data processing engine. It would seem to be possible to produce similar results using the various vendor systems and adjusting parameters. The workshop participants cited phantoms currently in use.

Dr. Franzblau wondered about a possible benefit in developing settings for different abnormalities. That idea, noted Dr. Siegel, suggests a benefit in performing processing at the workstation, varying the image. Prior to that, a unified "for processing" image could be helpful in, for example, relating to CT images. In any event, we will need guidance in comparing across vendor systems. Dr. Flynn agreed with that need but cited a difficulty in establishing settings among vendors. It would be helpful to obtain a set of images for a disease state. Dr. Henry proposed placing an indicator on the image (something imaged as the patient is imaged) to guide subsequent adjustment for display. The workshop participants considered this to be a good idea, although perhaps difficult to implement.

Ensuring Image Quality for Classification of Digital Chest Radiographs
– Ehsan Samei, Ph.D.

Dr. Samei noted downsides in the use of digital radiographs, including the following:

- Wide dynamic range can lead to over- or under-exposure of the patient
- Image post-processing can lack utility for physicians, reduce reading efficiency, and produce ad hoc images
- The digital format can lead to lost patient data and security problems

The potential advantages of digital radiography are not automatic. For full realization of digital radiography, users must recognize nuances associated with features, implementation, and quality control (QC). Quality control procedures can enable standardized processing and appearance and enable automated and optimal quantification. Metrics of image quality include resolution, noise, and signal-to-noise efficiency. Resolution is the ability to resolve distinct features, usually characterized by the Modulation Transfer Function (the efficiency of reproducing contrast at different spatial frequencies). It varies among common imaging systems. Noise refers to unwanted signals that interfere with interpretation and is best characterized by the Noise Power Spectrum (variance in terms of spatial frequencies). Signal-to-noise efficiency can be determined by the detector quantum efficiency. This too varies among the common digital radiography systems.

Dr. Samei described a possible quality control system for digital radiography that featured acceptance testing, system calibration, preventive maintenance, and periodic assessments. He also described the use of phantoms for quality control. He listed the following requirements for classifying pneumoconiosis:

- A robust QC program
- Standardized image acquisition protocols
- A consistent exposure index
- Raw image data in "for processing" form
- Consistent processing and display for consistent visualization across systems and cases
- Consistent analysis for automated quantification of pneumoconiosis
- Archives of raw and processed data for further analyses

Dr. Samei concluded that digital radiography can provide standardized classification of pneumoconiosis because of its quantitative nature and tractable performance characteristics. QC is essential to ensure robustness and integrity of data and to enable a reliable classification scheme. Dr. Samei recommended that NIOSH-affiliated programs enact maintenance and QC programs and follow predefined acquisition and processing protocols. NIOSH may consider maintaining a central Web server for affiliated facilities. Affiliated facilities could register their

imaging devices and performance metrics. NIOSH should consider accrediting affiliated facilities to ensure adherence to requirements.

Discussion

Dr. Alan Ducatman wondered whether phantoms offer consistency over time. Dr. Samei responded that facilities should be re-accredited following major changes. Dr. Vikas Kapil noted the problem of biases introduced in post-processing by readers. Dr. Samei suggested storing the raw images. Specific approaches to processing could be required for images to be classified. Dr. Krupinski suggested developing a program that could observe and save changes that are made during classifications—enabling future audits. Dr. Clunie suggested that there be quality control procedures for display systems. Dr. Flynn noted that the Modulation Transfer Function is not an issue in LCD systems with digital interfaces. He also wondered whether, because of the variety in acquisition devices, we should have a separate QC program for CR devices. Dr. Krupinski suggested that QC systems consider the reading environment (ambient lighting, etc.).

Standardizing File Formats, Security, and Integration of Digital Chest Image Files for Pneumoconiosis Classification – David A. Clunie, M.D., M.B.B.S.

Dr. David Clunie, of RadPharm Incorporated, noted that today there are no challengers to the DICOM standards for handling pneumoconiosis classification files. DICOM is supported by all modern devices in all countries. It has a bit depth suitable for the available sensors and features patient demographics, management information, and technique information in each header. It is the only inter-vendor standard in use. Dr. Clunie reviewed versions and features of DICOM, including limitations, and described DICOM as a system for thinking about interoperability. DICOM services include transfer across networks, querying for lists of patients and studies, retrieving studies, patients, series and images, creating work lists, and printing. Methods for transfer and workflow include the use of workstations, PACS, CDs, and networks.

Issues with CD viewers include the fact that images are often are burned to CDs with a viewer incorporated, the risk of transferring viruses on CDs, a need to be familiar with dozens of

viewers, a possible lack of grayscale pipeline support, and other concerns. One solution is to import standard media into a PACS. Barriers to importing include formats, ID reconciliation, and viruses. Software compatibility issues include multiple DICOM SOP classes, a need for "ready to view" images, and the need for a GSDF-calibrated display.

Image contrast features include a single default presentation of image contrast, a linear window center and width, and a nonlinear contrast adjustment. Dr. Clunie described the use of look-up table data. He stated that reference images are the ILO reference set and they can be displayed digitally with patient images. Displays are traditional PACS double portrait 3-megapixel workstations. Classifications can be performed using existing infrastructures and remotely, with images provided by a central server. In the future, authorized B readers might be able to access patient-related images and documents in large national databases. Dr. Clunie reviewed security issues surrounding patient images and other patient information, noting that digital data are at risk when in physical form (CD) or online. Privacy can be maintained by, for example, replacing a patient's name and social security number with a pseudonymous identifier.

Dr. Clunie described the DICOM structured reporting methods, which feature a variety of templates, and their advantages. He summarized his talk with the following statements:

- An entire infrastructure already exists to support clinical use of digital projection x rays
- It is based on the use of the current DICOM standard between modalities, PACS, and workstations, using networks and CDs
- Most sites are now experienced with exporting and providing outside access to digital images (including "for presentation" digital x-ray)
- The correct choice of an appropriate image viewer should allow consistent display and reliable review of images side by side with ILO or equivalent reference images
- Expensive displays already installed can easily be reused
- Results can be stored as DICOM Structured Reports—DICOM can support the addition of templates and codes.
- Matters of security and privacy can and should be addressed through conventional means that are already widely used clinically

Dr. Clunie recommended the following:

- Both CR and DX DICOM images should be permitted
- Processed "for presentation" images should be required, and they should not be dependent on proprietary processing in a display workstation
- Display workstations should be qualified and certified for use of ILO classifications by B readers, working with test images from different vendors and software, supporting variations of encoding and grayscale pipeline, and capable of displaying subject and reference images side-by-side
- For privacy, images should be de-identified before sending for reading
- A digital (not digitized film) reference set should be created and released, comparable in contrast and resolution to CR and DX images
- NIOSH should consider the creation of a managed distributed or centralized infrastructure, with remote reading and an open archive

Discussion

Dr. Prior wondered whether the use of de-identified data will be feasible. Dr. Clunie suggested that it will be, and the export of data will be for further patient care. He added that the approach for file handling that he described could be applied on an international level. Dr. Wagner wondered whether implementing such a system might be overly cumbersome. Dr. Clunie noted that many of the issues, such as multi-vendor PACS, have already been addressed.

Dr. Flynn expressed concerns about data export on CDs. Some large centers running full PACS operations are likely to export CDs in a proprietary format, thereby leading to a problem in reading them remotely. Dr. Clunie suggested that this problem may not be widespread. One solution would be for such a center to forward the data to a third-party CD writer that uses a non-proprietary DICOM format. Dr. Clunie raised another problem—some PACS systems alter the images (pixels, headers, etc.). Internet transmission engenders policy issues.

Dr. Clunie noted that other digital images (CT, MR) are used routinely in litigation, so that digital radiography likely will be used as well. Regarding privacy, a general recommendation is to obtain consent for secondary use of de-identified data up front. Dr. Prior added that, for matters such as privacy and litigation, it will be important to determine the form of data (that is, the image) that was observed by the reader. Dr. Clunie suggested that archiving requirements, backed by NIOSH, could reduce the possibility of altering images (as for malevolent purposes). Software updates should be accompanied by facility re-accreditation. Dr. Petsonk noted that the use of film-screen images will persist for some time, necessitating two tracks.

SMALL GROUP DISCUSSIONS AND RECOMMENDATIONS

The workshop participants divided into three smaller groups to discuss separately the following areas and provide recommendations to NIOSH and ILO for steps to be taken in the shorter term and longer term:

- Digital chest radiograph image acquisition and formation, including QC
- Image presentation, including processing and display
- File interchange, including formats and interoperability

They then reconvened as a whole to hear the leader of each subgroup report on the separate discussions.

1 – Digital Chest Radiograph Image Acquisition and Formation, Including QC

Dr. Samei summarized the results of the first subgroup's discussion of acquisition and formation of images. The group members agreed that implementation should be left to a follow-up initiative. They agreed on a multi-phase approach that would be grounded on the use of "for processing" data and basic image metrics. The consistency of image appearance must be a main goal. The subgroup described a three-phase approach:

- Within 1 year, establish guidelines and an approval process for facilities, equipment, and image quality. Seek consistency for the quality assurance processes with oversight and documentation. Define acquisition protocols with guidelines (scatter reduction, beam quality, rating of generators, exposure index, exposure monitoring over time, etc.). Define file format and QC guidelines. Conduct beta testing of the approach.

- In the medium term (1–3 years), create a more representative and complete phantom with automated analysis. Determine exposure requirements. Develop integrated QC with the phantom.

- In the long term (3–6 years), develop automated disease classification. Develop automated image quality assessment based on image data.

Discussion

Dr. Clunie wondered about the possibility of having influence over the sites' acquisition of equipment. We need to define values such as exposure index and influence vendors, perhaps through a NIOSH request for proposals. Dr. Samei cited the goal of classification in the requirements proposed. Dr. Ravin emphasized the basic point that digital images are better than film images. Dr. Henry noted a paradoxical trend in which radiographers produce poorer-quality images as a result of their overconfidence in a better technique (digital). Dr. Flynn suggested that exposure indicator/range may be a key value on which to focus. Dr. Weissman noted that NIOSH will have to engage Federal lawmakers to specify digital regulations or standards. A main goal, reminded Dr. Petsonk, is to protect the continuity of the program to protect the health of coal miners—a public health function. The workshop participants cited a need to establish an accreditation system or update the current NIOSH accreditation system.

2 – Image Presentation, Including Processing and Display

Dr. Flynn summarized the results of the second subgroup's discussion of image processing, focusing on three areas and making the following recommendations:

- Develop a reference case library for digital radiography to elaborate the effect of processing on classification, to provide a basis for a digital reference set, and to provide training material. Images would be in "for processing" form. Recent and ongoing studies could provide data for the library, which would be modest in size—perhaps a few thousand cases. Various institutions (e.g., National Library of Medicine, National Cancer Imaging Archive) have experience in developing such a collection. The NIOSH effort could include digital standard films, documentation on disease stages, and perhaps correlated film-screen and CT images.

- Develop consensus on display hardware specifications, for example, a 2-monitor, 24-inch, 3-megapixel system with grayscale cards, calibration, and navigation. This should be a discrete flat-panel technology, such as LCD. The system should include embedded test images, and there should be a document specifying aspects of the reading environment (ambient light, etc.).

- Develop display software for reading images and comparing them to standards. It should support a NIOSH/ILO viewer tailored to present reference comparisons, and it should be incorporated on a CD. One problem that must be avoided is the presence of an image from a prior reading left on a screen, which is read mistakenly as a current image. Cross-sectional cases present difficulties. A report template should be included in the electronic viewer presentation (it could then be easily transmitted electronically).

Discussion

Dr. Audrey Banyini noted that South Africa has a large database of images from medical surveillance, some of which might be shared with the reference case library. The workshop participants proposed standards and testing for reader vision. Dr. Henry wondered about incorporating the American Association of Physicists in Medicine test pattern in discs. Dr. Samei suggested that in addition NIOSH develop a reference image or pattern. The workshop participants agreed that there should be procedures for measuring and controlling, perhaps with

the aid of a light meter, the screen brightness and ambient lighting. The second subgroup proposed that, for image presentation in general, the first efforts involve tools that resemble film-screen tools (this could change over time).

3 – File Interchange, Including Formats and Interoperability

Dr. Clunie presented the results of the third subgroup's discussion of file interchange, listing assumptions, caveats, and requirements for image filing issues. The group assumed that the approach will be reusable in similar settings, will use NIOSH-supplied equipment, and will depend on a limited NIOSH support staff. There are questions about training and needed changes in regulations. The acquisition site must have pre-qualification of system and transfer processes. A central site will require pre-qualification processes and tools, QC processes and tools, and archival and disaster-recovery capabilities. Readers must receive images, read them, dispose of them, and send results.

Dr. Clunie listed the following recommendations for a CD-based solution in the short term (3-month):

- The acquisition site burns CR/DX to CD from a modality or PACS. Features include one patient per CD, identity in the header, a "for presentation" image (or, optimally, with "for processing" as well), the DICOM GP-CDR and IHE PDI profile, and no lossy compression. The site submits an initial pre-qualification CD.

- The central site receives and checks the CDs. Required tools include standalone PCs, displays, viewers, an automated CD format checker, an automated DICOM file checker, a CD duplicator, and a DICOM header editor. The central site duplicates CDs for archives and disaster recovery, sends CDs and ID documents to B readers, and receives the completed document from the B readers.

- The B readers' equipment is installed and calibrated by NIOSH and supplied in a configured state. The B readers complete reading and evaluation forms, then destroy the CDs.

- Viewer features include the following: custom or off-the-shelf, supporting calibrated 3 megapixel grayscale displays, read from a DICOM CD, a PA CXR display on left monitor, a single reference image on right monitor, scrolling through a reference set, supporting all grayscale variants, supporting window level and width with a sigmoid function, supporting pan and zoom, with identification and technique annotation and linear distance measurements.

Dr. Clunie listed the following recommendations for the longer term:

- The central site builds or uses an off-the-shelf PACS/RIS

- Acquisition sites submit CDs or send the information via the Internet

- Central site PACS match identifiers and archive images

- B readers view images on PACS remotely and securely over the Internet, complete forms, and make use of hardware as in the short-term solution

Dr. Clunie listed requirements for a reference set. The short-term recommended solution should assume the ILO 2008 standard, with highest-fidelity digitized data. DICOM encoding should feature DX "for presentation," contrast adjustment for P value grayscale output, window values either sigmoid or linear, black borders to reduce glare, identifying attributes in the header, spacing attributes that allow measurement of nodule size, orientation attributes that allow correct hanging, and validation as DICOM standard. The DICOM header identification should follow the current rules. The modality operator should be allowed to change social security number to a patient ID, to change date to DICOM study date, and to change the ALOSH approval number.

Facility name and address will be fixed by the field engineer. The central site (NIOSH) might have to perform some cleanup of headers during CD copying.

Discussion

It was noted that NIOSH will install B reader stations only for participants in the NIOSH program. As for a broader program, Dr. Clunie suggested that the paradigm for filing could be scaled-up. He noted that a header attribute could indicate no lossy compression. Responding to a question about their long term stability, he concluded that CDs might not be suitable for long-term archiving of data. Of course, the longer-term solution possibly will not use CDs at all. We should consider eventually a QC program for archiving.

THE ILO REVISION

Dr. Petsonk asked the workshop participants to offer ideas for revision of the ILO standards for reference images. How should the revision proceed? How should it reach out to partners? Should the ILO manufacture an image or accumulate and validate existing images?

Dr. Fedotov suggested that a next meeting focus more on international points of view as they relate to revising the guidelines. Tests of placing ILO standards on CDs have identified some problems. Perhaps NIOSH could work with ILO to prepare a CD with calibration and a test pattern and to conduct a study using various monitors. The result might lead to an international standard.

Dr. Balmes wondered whether the ILO was willing to develop generated standard images. In fact, noted Dr. Wagner, this is a research question. Dr. Samei stated that his research group has been performing simulations and validation and has learned that it takes about one year to develop a validated simulated image/pathology. Dr. Henry noted that actual candidate radiographs offer the benefit of cross-sectional comparison. Boundary radiographs of normal and abnormal cases would be helpful. Yet, noted Dr. Petsonk, how would we select from current

standard images? Dr. John Parker suggested considering only validated standard films, and he stated that we cannot simulate what dust does to the lung. Dr. Michael Jacobson stressed the need to maintain continuity with the past in advancing classification. Joseph Burkhart proposed a two-step approach featuring research followed by the acquisition of films.

平成 20 年度厚生労働科学研究費補助金（労働安全衛生総合研究事業）
分担研究報告書

じん肺分類密度の医師間変動を評価する新しい方法 ：CE-IRT モデル分析法の提案

研究協力者：松本　徹（放射線医学総合研究所）
分担研究者：日下幸則（福井大学医学部環境保健学）
分担研究者：志田寿夫（全国労働福祉協会）
：菅沼成文（高知大学医学部環境医学）
：田村太朗（福井大学医学部環境保健学）

研究要旨：
Item Response Theory ：IRT（項目反応理論）として一般に知られる項目（読影者）特性モデルに、我々が考案した Conflict-equilibrium：CE（葛藤—平衡）モデルを組み合わせたじん肺分類密度分析法(CE-IRT モデル分析法と命名)を提案した。本法の基本は 1980 年 ILO 実験データに示されている医師全体の読影傾向を ILO 標準写真と同じく国際的なじん肺分類傾向の標準と見做すことにある。これより統計変動の中に存在する、当該医師のじん肺分類密度と 1980ILO 標準の対応関係を、理論モデルに基づいて滑らかな曲線として顕在化し、医師個人が 0 型か 1 型以上か等 2 つのカテゴリ間の境界を分ける密度を特定する。次いで同一対象例のじん肺分類に参加した任意の二人の医師間のじん肺分類密度の変動および各医師のじん肺分類密度と 1980ILO 標準の偏りを定量する。ここでは提案手法の原理を示し、その妥当性を 1980 年版 ILO 標準写真選定実験データを対象に検証した結果を報告した。

A. 研究目的

アナログ画像である 1980ILO 標準写真[1]をデジタル X 線画像（CR,DR）の標準へ切り替えるための研究の一環として現在村田班は粉塵曝露歴が明らかなじん肺画像を班研究協力施設から収集している。筆者が想定するその後の手順は以下の通りである[2]。1)複数じん肺専門家が各施設から収集されたじん肺画像を鑑定し、当該画像のじん肺分類（密度、サイズ、形状他）を決定し、じん肺画像データベースを構築する。2)当該画像データベースからじん肺分類実験用画像を選択し、それを、我が国を代表する複数のじん肺診断医がじん肺分類する実験を行う。3)実験データを分析し、その結果に基づきデジタル標準画像候補を決定する。4)以下略。　本研究の目的は上記 3)のための分析法を提案し、その妥

当性を検証することである。

B. 研究方法

1. 本法が対象とするデータ

複数の画像を複数の医師が 1980ILO 標準画像を参照しつつ観察して各画像のじん肺密度を 12 階尺度分類する読影実験を想定する。その時、前節 3)の分析法に関与するデータを Fig.1 に示す。

ILO標準画像には標準密度-1②と表記したじん肺陰影が含まれるが、それらは、1980ILO 標準写真選定実験[3,4,5]で 27 人の医師が標準写真候補 106 枚をじん肺分類した密度「標準密度-0」①の一部であることに注目する。すなわち、ILO標準写真の密度は、じん肺密度分類用物差しの 1 型、2 型、3 型の mid-category(1/1,2/2,3/3)を表す目盛りに相当する。本法はこの大きな目盛りの間に、これまで誰も指摘しなかったので表立つことはなかったが細かい目盛り「標準密度-0①」が存在すると考え、これを 1980ILO 標準との偏りを定量分析するための物差しとして活用する。じん肺分類される画像は前節 1)で複数専門家によりじん肺分類結果が確定されたものでこれを「確定密度」③と表記する。1980ILO 標準写真とは別に、それぞれの国情によって定められた各国の標準写真の密度も ILO 標準ではないというレベルにおいて「確定密度」に属する。それらの確定密度写真を対象に各医師がじん肺分類した密度を「分類密度」④と表記する。①、②対③/④、③対④の関係(Fig.1 右図)から医師二人の分類密度④-1、④-2 間二次的関係 (Fig.1 左図) が得られる。

2. 本法が分析目標とするデータ間の関係

Fig.1①、②、③対二人の医師のじん肺分類密度④-1、④-2 の対応の基本形態とじん肺分類結果の期待値、非期待値の関係を Fig.2 に示す。医師二人のじん肺分類密度の対応(3)からその基になった(1)、(2)の一意的関係を復元することは原理的に不可能である。すなわち、(3)の関係のみからでは二人の医師が標準に照らして、確定密度または標準密度からどの程度偏ってじん肺分類を行ったか確認することはできない。同様に Fig.2(1)の結果から(3)は得られるが(2)は得られない。(2)から(1)、(3)の関係は必然の結果として明らかに出来る。本分析法は Fig.1(右図下辺)の細かい目盛り:「標準密度-0①」の物差しを使って、標準密度対分類密度/確定密度および二人の医師の分類密度間に存在する滑らかな関係 Fig.2(2)を後述する理論モデルに基づいて推測する。

3. 分析の流れ

本分析法全体の流れを Fig.3 に示す。以下これに沿って方法論を述べ適用結果を示す。

Fig.3　分析全体の流れ

4.　本分析法のじん肺密度分類の原理

4.1 葛藤—平衡（Conflict-Equilibrium：CE）モデル

読影者がじん肺 X 線画像を読影してじん肺陰影の有無、ここでは 0 型か 1 型以上か（以下これを 0 又は 1 と表現）判断する場面を想定する。その時 Fig.4 に示すような葛藤が心の中に生じ、葛藤の程度に応じた一定時間経過後、「1」という考えが心の中に占める割合（確信度：f_+）と「0」という考えの割合（確信度：f_-）が平衡状態に到達した時、読影者はじん肺陰影の有無（1,0）を決断すると仮定する。但し、$f_+ + f_- = 1$

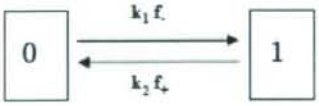

Fig.4　葛藤—平衡モデル

k_1 は「0」→「1」への遷移確率（「1」かもしれないと心が揺れる確率）、k_2 は「1」→「0」への遷移確率（「0」かもしれないと心が揺れる確率）を表す。

葛藤が平衡に達した時、

$k_1 f_- = k_2 f_+$

$k_1 (1-f_+) = k_2 f_+$

これより、

$f_+ = k_1 /(k_1 + k_2)$

ここで k_1、k_2 に対して以下を仮定する。

①k_2 ＝定数（読影が慎重なほど大きくなる。）

②$k_1 = k_0 \mathrm{EXP}[\alpha (S/N)]$

S は検出が期待されるターゲット（じん肺陰影）の信号成分、N はターゲットに関係した雑音成分、k_0 は比例定数、α は読影者（個人差：読影の癖、経験、学習）、画像(胸部 X 線写真、CT：撮影条件、表示条件)等により異なる係数

k_1 は S/N が大きくなると指数的に増大する。以上のことより、t= S/N とおくと、人が X 線画像を読影して、じん肺陰影があるという考えが心の中を占める割合（確信度）f_+は次式で表せる。

$$f_+(t)=k_0 \mathrm{EXP}(\alpha t)/[k_2 + k_0 \mathrm{EXP}(\alpha t)]$$
$$= 1/[1 + (k_2/k_0) \mathrm{EXP}(-\alpha t)]$$
$$=1/[1 + \mathrm{EXP}(-(\alpha t - a^*)] \qquad (1)$$

但し、$a^*=\ln (k_2 / k_0)$

(1)式で表される確信度 f_+ は単調増加関数であり、じん肺陰影密度 t が大きいほど高くなり、「1」と判断される確率が増大することを表す。

4.2 項目反応理論（Item Response Theory：IRT）モデル[6)]

じん肺密度分類を 0 か 1 かで行った時の IRT モデルとして(2)式の 2 パラメータロジスティックモデルを適用した。

$$P_j(\theta)=1/[1 + \exp\{-Da_j(\theta - b_j)\}]$$

$-\infty < \theta < \infty$ (2)

ここでjはIRTにおける項目番号を表す。本報では読影者を項目j、症例(画像)を被じん肺分類潜在特性値 θ (じん肺と判定されやすさの指数)と見立てた分析を行う。a_j, b_jは項目j (読影者) の特徴を表すパラメータで、aは(2)式を観測値の0,1データに当てはめた時の項目特性関数の傾斜を表す。bはその関数全体が特性値 θ 軸上、より右側にあるか、左側にあるかを表すパラメータである。$P_j(\theta)$は潜在特性値 θ における、読影者 (項目) jの「1」判断率を表す。(2)で表される関数は θ の単調増加関数であり、特性値 θ が大きいほど項目 (読影者) が「1」判断する確率 $P_j(\theta)$が増大することを表す。

4.3 CEモデルとIRTモデルの結合

CE モデルと IRT モデルは(1)式および(2)式において、$\alpha t - a^* = Da(\theta - b)$ と置くことにより対応付けられる。すなわち、定数；D=1.7、$a^* = -Dab$、$\alpha = Da$、$t = \beta\theta$: β は比例定数 (3)

$$f_+(\theta) = 1/[1 + EXP(-Da(\theta - b))] = P(\theta) \quad (4)$$

$$f_-(\theta) = 1 - f_+(\theta) = 1 - P(\theta)$$

後述する実験データの実際の分析では、じん肺陰影の有無判断 (1,0) の結果に IRT モデルの(2)式を当てはめて a, b, θ を求め、$P(\theta)$を計算する。CEモデルはじん肺陰影密度分類の問題において(2)式を用いることの理由やその意義を読影者の心理学的側面から説明し、変数；t を(3)、(4)により IRT モデルと結合し、θ、$P(\theta)$と関係付ける。

ここで改めて以上を整理すると、2つの理論では以下を仮定している。

①CE モデルにおいてじん肺陰影の存在確信度$f_+(\theta)$または$f_-(\theta)$の根拠となる、定量的、客観的パラメータ t=S/N=肺内ダスト量∝読影者により判定されるじん肺密度(1-12)∝0,1判定結果の存在および IRT モデルにおいてそれに対応する潜在特性 θ の存在

②CE モデルにおいて確信度特性曲線$f_+(\theta)$または$f_-(\theta)$の存在およびIRTモデルにおいてそれに対応する「1」判断率$P(\theta)$の存在

5. 実験材料

CE-IRT 法の有用性を検証するため以下の ILO 読影実験データを使用した。ILO は1977 年から約 2 年かけて標準写真選定のため国際的な読影実験を行った[3]。日本、米国、欧州から 106 枚の標準候補写真が収集され、これらの写真のコピー各1セットが各国代表に送付され、各国の読影者に回覧された。読影者は 27 人、内訳は、日本 9 人、米国 8 人、欧州 10 人であった。じん肺密度分類は 1971 年版 ILO 標準写真を参照して行われた。読影結果は ILO 作成のレポートシートに記録され、読影後シートは ILO 本部に返送された。これより、ILO は各写真に対する 27 人の読影者のじん肺診断結果を Fig.5 のようにまとめた。本報で対象としたデータは、読影者が 12 階尺度で回答した 3 種のじん肺密度分類結