

FAILURE TO IMPROVE OBSERVER PERFORMANCE WITH SCAN SMOOTHING

David E. Kuhl, Toby D. Sanders, Roy Q. Edwards, and P. Todd Makler, Jr.

Hospital of the University of Pennsylvania, Philadelphia, Pennsylvania

Data smoothing has attracted attention as a potential method for suppressing the appearance of noise in a scan picture so that important threshold features may become more apparent to the physician observer. One method is data bounding where statistically aberrant elements are replaced with values which are more likely to represent the real activity distribution (1-3). In this digital process, the count value of each picture element is compared to an average of its neighbors and replaced if it is beyond certain limits. Another method is spatial averaging which reduces the wide fluctuation between elements resulting from poor counting statistics. The count value of each element is replaced by an average of the values in neighboring elements either by shaping the light spot in photorecording (4-10) or with secondary digital processing (1-3,11-15). Spatial averaging is in more common use than data bounding.

The change in appearance of a picture smoothed with spatial averaging depends on the shape and spatial extent of the averaging function. If detection depended on signal-to-noise ratio alone in a threshold situation where the input signal was very near the background count, the optimum averaging function should equal the intensity distribution of the input signal (matched filter) (14,16,17). But an averaging function of such large spatial extent is known to degrade spatial resolution adversely and is seldom used. Difference of opinion as to optimum averaging function has resulted in suggested strategies for variable processing where the averaging function may be chosen to fit the picture characteristics (2,3,13,18). But it is not yet clear how optimum choices for smoothing are to be made and what improvement in physician performance can be expected.

Physician performance must be the real test of worth here. When the human observer is considered, the conditions of viewing and the character of display become important. For optimum viewing conditions, the visual angle subtended by the target

should be carefully chosen, yet viewing distance is frequently neglected in tests of scan smoothing. Morgan (16,19) has explained the importance of generous viewing distance for scan interpretation. Scan images characteristically have low spatial frequency components, and long viewing distances or minification are required if there is to be an optimum match to the maximum retinal response. The more narrow visual angle also will cause the eye to respond more poorly to higher frequency components of picture noise, a major goal of processing with spatial averaging (20).

In this project, we sought to clarify the practical value of bounding and spatial averaging applied to phantom and brain scan pictures of near threshold lesions viewed by observers under near optimum conditions with results analyzed according to signal detection theory.

PHANTOM SCAN EXPERIMENT

Method. The phantom was a radioactive (^{99m}Tc) sphere (object) measuring 2.4 cm in diam superimposed upon a uniform disk of radioactivity (background) which in turn was surrounded by a narrow annulus approximately twice as radioactive as the disk (Fig. 1). A scan picture was a 50×50 matrix of picture elements, each measuring 0.375×0.375 cm. The object was at the focal length of the collimator* where the FWHM was 1.6 cm. The picture element size was sufficiently small compared with the width of collimator line spread function to avoid any significant resolution loss due to digital sampling.

The basic set of original data (Process 1) was nine scans with the object and 12 without, divided equally

Received Oct. 20, 1971; revision accepted May 12, 1972.

For reprints contact: David E. Kuhl, Div. of Nuclear Medicine, Dept. of Radiology, Hospital of the University of Pennsylvania, 3400 Spruce St., Philadelphia, Pa. 19104.

* Picker 19-hole collimator, No. 2107A.

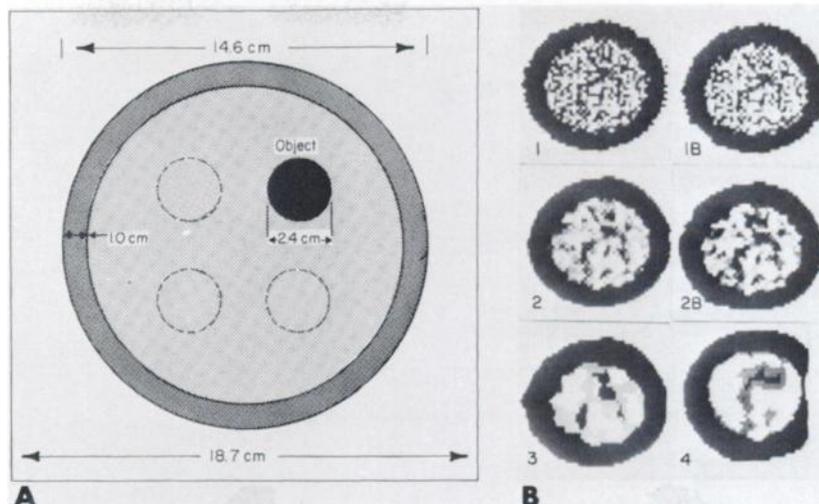


FIG. 1. (A) Phantom configuration. Object could be present in any one of four locations or could be absent. (B) Example of original picture and each of five pictures processed from it, as photographed from CRT screen. Observer controlled contrast and off-set during interpretation and was not restricted to number of grey steps shown in photographs.

among backgrounds of 50, 100, or 300 counts/cm². In scans with an object, there was a signal-to-noise constant* equal to 1.8 ± 0.3 (near threshold for visual perception).

The hybrid processing and display system used in these experiments is described in detail elsewhere (3). We recorded the original scan data on perforated paper tapes, and then processed them in our subsystem computer† producing other perforated tapes representing secondary data pictures. We processed each scan of original data five different ways using data bounding or spatial averaging.

In data bounding (3) the process permitted 25% count excursion for each picture element compared to the mean of the counts in eight surrounding elements. Otherwise, the count value of each picture element was replaced by the high or low limit. For spatial averaging (3) the process replaced the count value of each picture element by a weighted average of surrounding elements plus the center element. The desired spread functions were obtained by appropriate weighting and reprocessing of the secondary tapes. Process 1 was the original data; Processes 2–4 gave increasing smoothing areas as shown in Fig. 2. Bounding was performed both on the original data alone (Process 1B) and prior to minimal spatial averaging (Process 2B).

Comparing the components of this test system, then, the object intensity distribution nearly equals the spread function of the collimator. The input signal is the convolution of the two and nearly equals

the processor spread function of Process 4, which serves in this instance as a matched filter.

Processing produced a complete set of 126 tapes, each representing one test picture (six variations for nine pictures with object, and six variations for 12 pictures without). The test tapes were randomized and then, in sets of eight, were loaded onto a magnetic storage drum. Each picture was available for presentation to the observer in sequence by pushing a selector button. An entire scan raster was recycled 30 times/sec, presented repetitively to the cathode-ray tube* (CRT), and the picture appeared on the screen as a persistent image with very little perceptible flicker. Since the object was positioned off-center in the picture, the test supervisor could cause it to have any one of four randomized positions by inverting or reversing the presentation on the CRT.

Contrast and offset were under the direct control of the observer at the time of viewing. Each observer was able to adjust each picture, under observation, to any contrast or offset, including a gradual change from white to black between any regions of interest (Fig. 1). In this way, the observer could take maximum advantage of the 6-bit count content of each picture element and could rapidly examine data in high- or low-count regions without the limitations imposed when one is restricted to the number of grey steps in a single static picture on a CRT screen.

The optimum visual angle was chosen according to the data of Morgan (19) based on an object diameter of 2.4 cm displayed on the CRT screen with a 3:1 minification. We achieved the required viewing distance, 250 cm, using a mirror system that per-

* Signal-to-noise constant is ratio of difference between object and background counts to root-mean-square variation of background counts, where both object and background counts are summed within equivalent disks (4 cm diam).

† DATA 620/i Varian Data Machines, Irvine, Calif.

* EIA Model 10ALP7 (yellow layer only) phosphor with amber filter. Screen size: 8 × 10 cm. Maximum luminance: 69 cd/meter². Minimum luminance: 4.5 cd/meter². Ambient illumination: 2.7 lx.

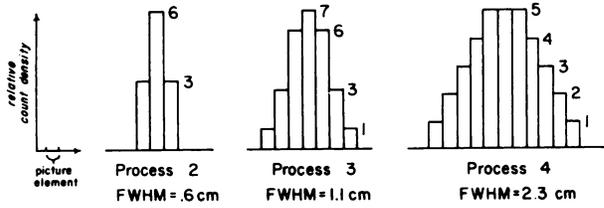


FIG. 2. Spatial averaging (minimal, moderate, maximum). Figures are line-by-line summations of two-dimensional distribution of relative count density that would result if each process were applied to single picture element of data.

mitted the observer to see the screen at this fixed distance and yet to maintain access to the front panel controls.

Eight observers examined each of the 126 pictures in turn, decided whether or not an object image was present on the picture, and if so, determined in which quadrant it was located. Each observer had been told that a single barely visible object was present on some of the pictures. Viewing time was not limited, but no observer spent more than 20 sec examining any one picture. A correct response on positive pictures required correct detection and localization.

Results. The means of the true positive rates $[P_{SN}(A)]$ and the false positive rates $[P_N(A)]$ for the observers are shown in Fig. 3. Trend of change in the true positive rate is generally followed by a parallel trend in the false positive rate. Both $[P_{SN}(A)]$ and $[P_N(A)]$ had wide interobserver variations due mainly to differences in criteria for decision. Some observers were strict in deciding when

the picture was a positive one; some were lax. Those who were strict would have a low $[P_N(A)]$ and also a low $[P_{SN}(A)]$, whereas with those who were lax, both were increased.

The relationships between true and false positive rates have been described mathematically in signal detection theory (21-23). This defines an index of detectability, d' , which is useful as a measure of observer performance*. When $[P_{SN}(A)]$ and $[P_N(A)]$ are known, the value of d' can be read from a table (22). The index of detectability, d' , depends on the observer's sensory ability to detect a signal, not on the observer's criteria for judgment. If the processing of the pictures in this experiment increased the sensitivity or performance of the observers, there should be an increase in d' . Figure 4 is a plot of the mean of observers d' in this experiment against the entire range of d' that is possible. In the group as a whole, no trend of improvement is apparent throughout the series of processes.

We then tested for possible improvement in performance if we assumed nearly equal sensory ability among the observers. For each observer, we calculated d^* , the deviation of his d' for each process from

* This index of detectability determines the separation of probability density functions of detection and false alarm rate and incorporates a comparison of obtained performance with ideal performance. The measure d' may be derived from a receiver operating characteristic (ROC) curve, that is, a plot of true positive rate compared with false positive rate on normal-normal coordinates. In this paper, however, the values for d' were obtained from tables (22) using the experimentally determined values for true and false positive rates.

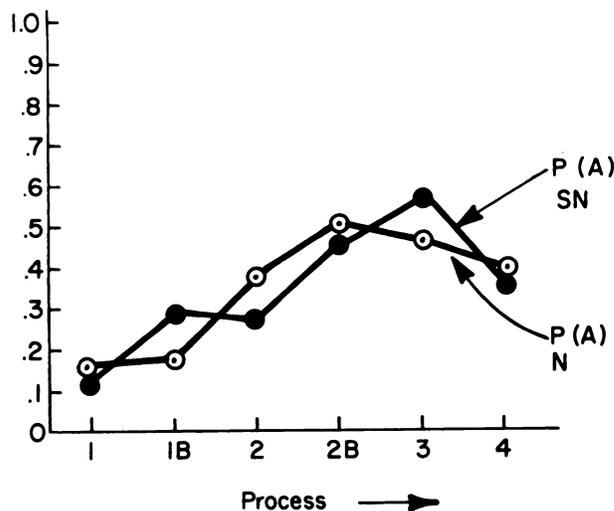


FIG. 3. Phantom scan experiment. Processing resulted in increase in both mean true positive rate $[P_{SN}(A)]$ and mean false positive rate $[P_N(A)]$. If choices had been made by chance alone, we would expect $[P_{SN}(A)] = 0.125$ and $[P_N(A)] = 0.5$.

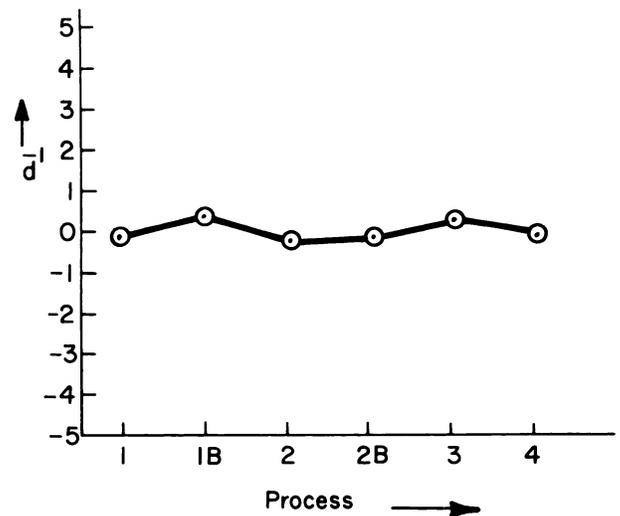


FIG. 4. Phantom scan experiment. Mean of observer's index of detectability, d' , does not increase (improve) with processing. If choices had been made by chance alone, we would expect $d' = -1.15$.

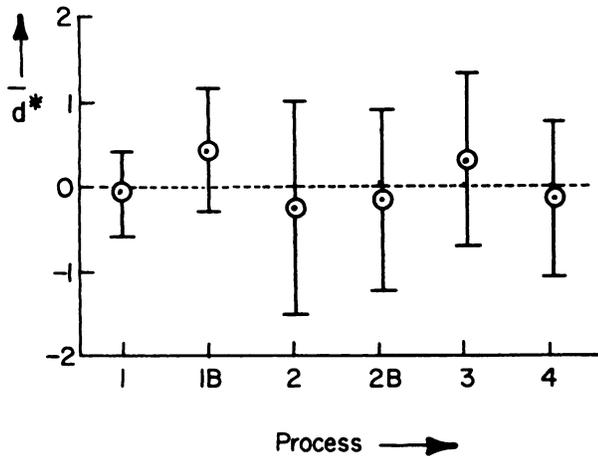


FIG. 5. Phantom scan experiment. Limits of mean d^* among observers include zero for all processes. This indicates none of processes were improvement over original data.

his mean d' for all processes. We then determined the mean d^* among observers for each process†. These values are shown in Fig. 5 along with 95% confidence limits. Since all 95% confidence intervals include zero, there is no difference in detectability d' , that is, none of the processes were an improvement over the original data.

BRAIN SCAN EXPERIMENT

Does minimal smoothing for cosmetic alteration of a brain scan picture improve observer performance?

Although we found no advantage to applying these smoothing processes to low contrast phantom scan data, we considered the possibility of improvement in the more realistic situation of brain scan interpretation. Here, experienced observers are required not only to detect and localize but to recognize as well, that is, to distinguish between normal and abnormal structures. In this experiment only a single process was compared to the original data. Given a choice among the smoothing processes as applied to brain scans, three experienced observers preferred Process 2B, bounding followed by minimal spatial averaging, rather than more extensive smoothing because it seemed to give the best compromise of noise reduction and artifact.

Method. The test scans selected were 88 pictures composed of four-view studies of 11 normal patients

† For each observer:

$$\bar{d}' = \frac{d'_{1} + d'_{1B} + d'_{2} + d'_{2B} + d'_{3} + d'_{4}}{6}$$

$$d^* = d' - \bar{d}'$$

Therefore, we can calculate a mean d^* for the eight observers.

and 11 patients with positive scans. At the conclusion of the observer test and with knowledge of lesion position in each picture of the abnormal series, we scored lesion visibility on a scale from 0 to +4, where +4 was the most obvious: 13 lesions were +4, 11 were +3, 7 were +2, 6 were +1, and 7 were 0 (Fig. 6). The background count in the center of each head image averaged 250 counts/cm².

The three observers were physicians experienced in scan reading. The display conditions were the same as those described in the phantom experiment. First the observers scored randomized pictures of Process 1 (original data) and Process 2B separately, and then in pairs. When viewed in pairs for comparison, the observer alternated a Process 1 picture with a Process 2B picture by throwing a panel switch.

Results. The results shown in Figs. 7 and 8 are similar to those of the phantom experiments. An increase in the true positive rate with processing was accompanied by an increase in the false positive rate. The index of detectability, d' , was unchanged whether the processed scan was observed alone or together with the original data.

CONCLUSIONS

In principle, picture minification and picture smoothing have potential for improving observer

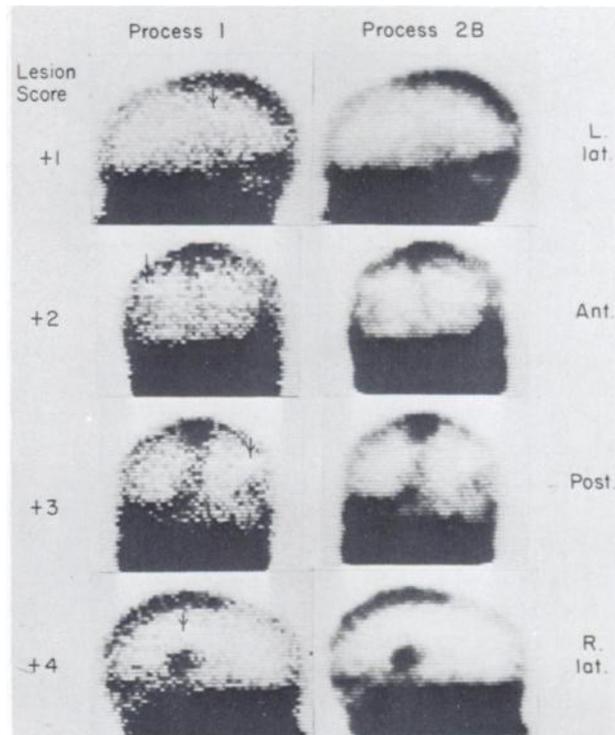


FIG. 6. Brain scan experiment. Example of lesion scoring based on prior knowledge of lesion position.

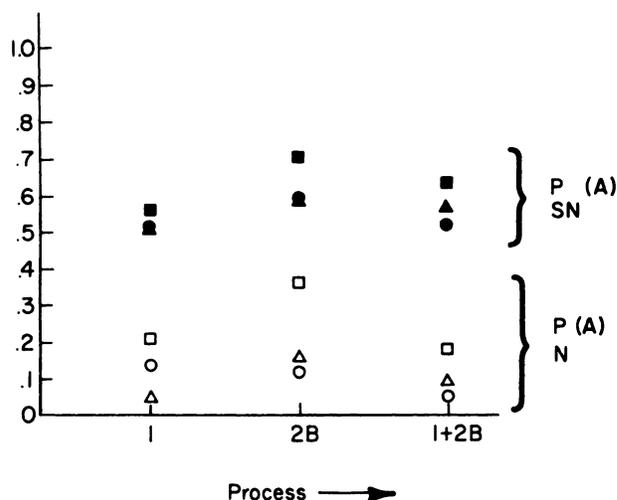


FIG. 7. Brain scan experiment. Process 2B increased true positive rate [$P_{SN}(A)$] and false positive rate [$P_N(A)$]. Results from reading processed and unprocessed pictures in sequence (Processes 1 + 2B) were no different than reading unprocessed pictures alone.

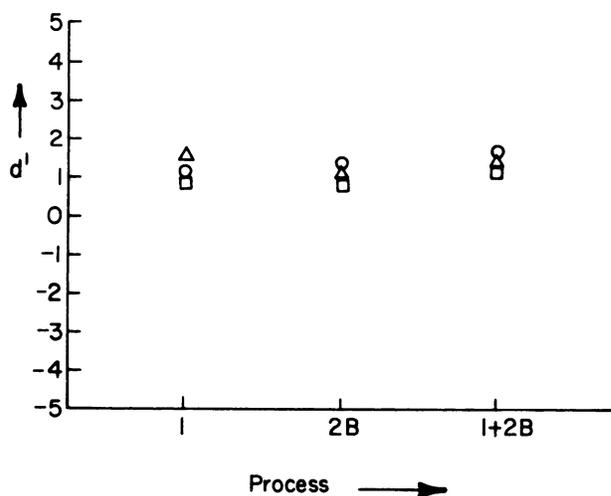


FIG. 8. Brain scan experiment. Index of detectability, d' , is unchanged from reading unprocessed pictures (Process 1) whether processed pictures were observed alone (Process 2B) or together with original data (Processes 1 + 2B).

performance in detecting threshold objects. It was not apparent what additional advantage was offered by smoothing after prior minification. In these experiments, we failed to show any improvement of observer performance in detecting low contrast images when they were viewed at a distance appropriate to match the most significant components of the spatial frequency spectrum of the object to the optimum spatial frequency response of the eye. Although smoothing did increase true positive rate, it also increased false positive rate as quantum mottle from insufficient statistics blended into spurious structures.

Morgan (19) has estimated that optimum viewing

distance is approximately 340 times the diameter of the object in view. Scans are usually read at much shorter distances. Scans could easily be interpreted under more optimum conditions with appropriate picture minification and the use of a diminishing lens. Our study suggests that in the usual clinical situation this approach is a more useful investment in effort than picture processing with data bounding or spatial averaging as we have applied it.

ACKNOWLEDGMENT

We thank Anthony R. Ricci, Ellen B. Gallagher, Theodore P. Sanders, Robert W. Burt, Michael J. Eymontt, and Dan H. Moore for assistance. USAEC Contract AT(30-1)-3175, USPHS Research Grant GM-16248, USPHS Research Career Program Award CA-14020 (DEK), ACS Institutional Grant IN38-L, Sub-project No. 19 (PTM), and a James Picker Medical Student Scholarship (PTM) supported this project. AEC identification No. NYO-3175-68.

REFERENCES

1. BROWN DW: Digital computer analysis and display of the radioisotope scan. *J Nucl Med* 5: 802-806, 1964
2. KUHL DE, EDWARDS RQ: Digital techniques for on-site scan data processing. In *Fundamental Problems in Scanning*, Gottschalk A, Beck R, eds, Springfield, Ill, CC Thomas, 1968
3. KUHL DE, EDWARDS RQ: A hybrid processor for modifying and rearranging radionuclide scan data under direct observation. *Radiology* 92: 558-570, 1969
4. MACINTYRE WJ, CHRISTIE JH: The use of data blending to reduce statistical fluctuations in radioisotope scanning. *Radiology* 86: 141-142, 1966
5. MACINTYRE WJ, CHRISTIE JH: A comparison of data averaging of radioisotope scan data by photographic and dimensional computer techniques. In *Medical Radioisotope Scintigraphy*, vol 1, Vienna, IAEA, 1969, pp 771-782
6. BECK RN, CHARLESTON DB, EIDELBERG P, et al: The ACRH brain scanning system. *J Nucl Med* 8: 1-24, 1967
7. FREEDMAN GS, WOLBERG JR, JOHNSON PM: Comparison of the information content of data-blended and conventional photoscans. *Radiology* 90: 921-924, 1968
8. ATKINS HL, HAUSER W, RICHARDS P: A comparison of conventional and blended display technics in scintiscanning. *Radiology* 90: 912-920, 1968
9. ATKINS HL, HAUSER W, RICHARDS P, et al: Analysis of the optimum conditions for data display and image perception in photoscanning. In *Medical Radioisotope Scintigraphy*, vol 1, Vienna, IAEA, 1969, pp 759-770
10. WYPER DJ: A comparison of blended and conventional photoscan displays. *Radiology* 99: 147-151, 1971
11. SCHEPERS II, WINKLER C: An automatic scanning system using a tape perforator and computer techniques. In *Medical Radioisotope Scanning*, vol 1, Vienna, IAEA, 1964, pp 321-329
12. SPRAU AC, TAUXE WN, CHAAPEL DW: A computerized radioisotope scan-data filter based on a system response to point source. *Proc Staff Meet Mayo Clin* 41: 585-598, 1966
13. PIZER SM, VETTER HG: The problem of display in the visualization of radioisotope distributions. *J Nucl Med* 7: 773-780, 1966

14. TANAKA E, IINUMA TA: Approaches to optimal data processing in radioisotope imaging. *Phys Med Biol* 15: 683-694, 1970

15. MAROGLIO LJ, KNOWLES LG, KOHLENSTEIN LC, et al: Improvement of lesion detection performance with image processing. *J Nucl Med* 12: 381-382, 1971

16. MORGAN RH: Concluding remarks. In *Fundamental Problems in Scanning*, Gottschalk A, Beck R, eds, Springfield, Ill, CC Thomas, 1968, pp 385-390

17. ROSSMAN K: Detail visibility in the presence of quantum fluctuations. In *Fundamental Problems in Scanning*, Gottschalk A, Beck R, eds, Springfield, Ill, CC Thomas, 1968, pp 326-335

18. ANGER HO, VAN DYKE DS, GOTTSCHALK A, et al:

The scintillation camera in diagnosis and research. *Nucleonics* 23, No. 1: 57, 1965

19. MORGAN RH: Visual perception in fluoroscopy and radiography. *Radiology* 86: 403-416, 1966

20. GREGG EC: Modulation transfer function, information capacity and performance criteria of scintiscans. *J Nucl Med* 9: 116-127, 1968

21. GREEN DM, SWETS JA: *Signal Detection Theory and Psychophysics*. New York, John Wiley, 1966

22. ELLIOTT PB: Tables of d' . In *Signal Detection and Recognition by Human Observers*, Swets JA, ed, New York, John Wiley, 1966, pp 651-684

23. KUNDEL HL, REVESZ G, SHEA FJ: Display format and decision-making in television processing of chest radiographs. *Invest Radiol* 4: 264-268, 1969

ACCEPTED ARTICLES

TO APPEAR IN UPCOMING ISSUES

A Study of Normal Mammary Lymphatic Drainage by means of Radioactive Isotopes. Accepted 5/11/72.

E. Vendrell-Torné, J. Setoain-Quinquer, and F. M. Domenech-Torné

Radionuclides in Renal Transplantation. Accepted 5/11/72.

G. Hor, H. W. Pabst, K. J. Pfeifer, P. Heidenreich, H. Langhammer, and H. G. Heinze

Abnormal Liver Scan (Radiocolloid "Hot" Spot) Associated with Superior Vena Caval Obstruction (Case Report). Accepted 5/26/72.

John T. Joyner

Uptake of Radiostrotrium in Lungs and Other Extrasosseous Tissues (Letter to the Editor). Accepted 6/8/72.

Tapan K. Chaudhuri, Tuhin K. Chaudhuri, and James H. Christie

Splenic Sequestration of ^{99m}Tc -Labeled Red Blood Cells. Accepted 6/8/72.

Harold L. Atkins, William C. Eckelman, Wolfgang Hauser, Johannes F. Klopper, and Powell Richards

Determination of Plasma Renin Activity by Radioimmunoassay: Comparison of Results from Two Commercial Kits with Bioassay. Accepted 6/8/72.

L. Rao Chervu, Marc Lory, Theresa Liang, Hyo Bok Lee, and M. Donald Blaufox

Letter to the Editor. Accepted 6/12/72.

Robert R. Hiscock

Quantitative Radioisotopic Angiocardiology. Accepted 6/12/72.

Paul M. Weber, Leon V. dos Remedios, and Ivan A. Jasko

Renal Clearance and Brain Tumor Localization in Mice of ^{99m}Tc Compounds of (Sn) DTPA, (Iron-Ascorbic Acid) DTPA, and Iron-Ascorbic Acid. Accepted 6/26/72.

Ted Konikowski, Howard J. Glenn, and Thomas P. Haynie

New Bone Scanning Agent: ^{99m}Tc -Labeled 1-Hydroxy-1, 1-Disodium Phosphonate. Accepted 6/26/72.

Frank P. Castronovo and Ronald J. Callahan

Effect of Increased Iodine Intake on Thyroidal Response to TSH Stimulation. Accepted 6/26/72.

Jerald C. Nelson

Brain Scanning with $^{99m}\text{TcO}_4^-$ in Multiple Sclerosis (Concise Communication). Accepted 6/26/72.

David C. Moses, Larry E. Davis, and Henry N. Wagner, Jr.

Rapid Preparation of Autologous Radioiodinated Fibrinogen. (Concise Communication). Accepted 6/26/72.

Ronald C. Roberts, Carolyn O. Sonnentag, and James H. Frisbie

Optimization of Data in Lung and Brain Scans. Accepted 6/26/72.

Suresh M. Brahmavar, Robert B. Groves, Edward V. Staab, Jon J. Erickson, and A. Bertrand Brill

Graves' Disease with Functioning Nodules (Marine-Lenhart Syndrome). Accepted 6/26/72.

N. David Charkes

Serum Tri Iodothyronine Uptake Using Albumin Microspheres in the Assessment of Thyroid Function. Accepted 6/27/72.

E. Rollerli, G. Buzzigoli, G. Plassio

Filter to Correct for Inverse Square Law Nonuniformity in the Pin-hole Collimator (Letter to the Editor). Accepted 6/28/72.

Herbert Malamud

Pulmonary Metastases from Thyroid Carcinoma Detectable only by ^{131}I Scan. Treatment and Response (Case Report). Accepted 6/28/72.

J. E. Turner and G. J. Weir, Jr.

Macroaggregation of an Albumin-Stabilized Technetium-Tin (II) Colloid (Concise Communication). Accepted 6/28/72.

Max S. Lin and H. Saul Winchell

Whole Body Imaging and Count Profiling with a Modified Anger Camera. I. Principles and Application. Accepted 6/28/72.

Michael Cooke and Ervin Kaplan

Whole Body Imaging and Count Profiling with a Modified Anger Camera. II. Implementation and Evaluation. Accepted 6/28/72.

Michael Cooke and Ervin Kaplan

Average Geometrical Value in Absorbed Dose Calculation (Letter to the Editor). Accepted 6/29/72.

Richard G. Lane

Protection Against Radiation from Brachytherapy Services (Book Review). Accepted 6/29/72.

Gerald J. Hine

Frontiers of Nuclear Medicine (Aktuelle Nuklearmedizin) (Book Review). Accepted 6/29/72.

Ngo Tran

Splenic Displacement and Changes of Splenic Size (Letter to the Editor). Accepted 7/6/72.

Bruno Schober

Reply to the Letter to the Editor. Accepted 7/6/72.

Richard P. Spencer

^{99m}Tc -EHDP: A potential radiopharmaceutical for Skeletal Imaging (Preliminary Note). Accepted 7/7/72.

G. Subramanian, J. G. McAfee, R. J. Blair, E. G. Bell, A. Mehter, and T. Connor

Hepatic Scintigraphy in Clinical Decision Making. Accepted 7/7/72.

David E. Drum and Julia Sdougou Christacopoulos

Regional Cerebral Blood Flow with the Anger Camera. Accepted 7/20/72.

B. Leonard Holman, Rex Hill, David O. Davis, and E. James Potchen

Accelerated Clearance of Radioactive Chelate from Cerebrospinal Fluid in Experimental Meningitis. Accepted 7/20/72.

Prantika Som, Fazle Hosain and Henry N. Wagner, Jr.

False-Positive Liver Scan Due to Lung Abscess (Case Report). Accepted 7/20/72.

Harry K. Genant and Paul B. Hoffer

Gating Mechanism for Motion-Free Liver and Lung Scintigraphy (Concise Communication). Accepted 7/20/72.

Frank H. DeLand and Walter Mauderli

Diagnostic and Therapeutic Implications of Long Term Radioisotope Scanning in the Management of Thyroid Cancer. Accepted 7/27/72.

G. T. Krishnamurthy and W. H. Biah