

# Quantification of SPECT Concentric Ring Artifacts by Radiomics and Radial Features

**Emilio Mezzenga**

IRST: Istituto Scientifico Romagnolo per lo Studio e la Cura dei Tumori Srl

**Anna Samelli**

IRST: Istituto Scientifico Romagnolo per lo Studio e la Cura dei Tumori Srl

**Giovanni Bellomo**

IRST: Istituto Scientifico Romagnolo per lo Studio e la Cura dei Tumori Srl

**Frank P DiFilippo**

Cleveland Clinic

**Christopher J Palestro**

Donald and Barbara Zucker School of Medicine at Hofstra/Northwell

**Kenneth J Nichols** (✉ [knichols@northwell.edu](mailto:knichols@northwell.edu))

Northwell Health <https://orcid.org/0000-0003-2010-7078>

---

## Original research

**Keywords:** Image artifacts, SPECT, gamma camera, quality assurance, image texture analysis, inter-observer agreement, phantom

**Posted Date:** September 18th, 2020

**DOI:** <https://doi.org/10.21203/rs.3.rs-78752/v1>

**License:**   This work is licensed under a Creative Commons Attribution 4.0 International License.

[Read Full License](#)

---

1           **Quantification of SPECT concentric ring artifacts by radiomics and radial features**

2  
3           **Emilio Mezzenga,<sup>1</sup> Anna Sarnelli,<sup>1</sup> Gioavanni Bellomo,<sup>2</sup> Frank P. DiFilippo,<sup>3</sup> Christopher**  
4           **J. Palestro,<sup>4</sup> Kenneth J. Nichols<sup>4</sup>**

5  
6           <sup>1</sup>Medical Physics Unit, Istituto Scientifico Romagnolo per lo Studio e la Cura dei Tumori (IRST)  
7           IRCCS, Meldola, Forli-Cesena, Italy

8           <sup>2</sup>Neurology clinic, Department of Medicine, University of Perugia, Perugia, 06132 Italy

9           <sup>3</sup>Cleveland Clinic, Department of Nuclear Medicine, Cleveland, Ohio, 44195 USA

10          <sup>4</sup>Department of Radiology, Donald and Barbara Zucker School of Medicine at  
11          Hofstra/Northwell, Hempstead, New York, 11549 USA

12  
13          Please address all correspondence to: Kenneth J. Nichols, PhD; Division of Nuclear Medicine &  
14          Molecular Imaging; Northwell Health; 270-05 76<sup>th</sup> Avenue; New Hyde Park, NY 11040 USA;  
15          Electronic mail:[knichols@northwell.edu](mailto:knichols@northwell.edu)

16  
17

18

19

20

21

22

23 **Abstract**

24 **Background:** Conspicuous concentric ring artifacts in phantom reconstructions triggers retuning  
25 SPECT systems. These evaluations are visual, not quantitative. Our study was undertaken to  
26 determine the degree to which observers agree about SPECT concentric ring artifacts, and to test  
27 whether quantitative texture analysis metrics correspond to significant artifacts.

28 **Methods:** Test data were acquired as part of quarterly quality assurance using standardized  
29 SPECT phantoms containing solid spheres, solid rods and volumes of uniform activity  
30 concentration loaded with  $^{99m}\text{Tc}$ . Forty SPECT studies were identified as having concentric ring  
31 artifacts or were acquired to assess whether artifacts were resolved following camera retuning  
32 after obtaining an unacceptably non-uniform result. Transaxial reconstructions were reviewed  
33 independently by two medical physicists who graded severity of artifacts on a 5-point scale.  
34 Counts were tabulated in volumes of interest created in uniform phantom sections, from which  
35 were computed 72 radiomics image texture analysis metrics. Radial contrast ( $R_{\text{Contrast}}$ ) derived  
36 from the radial profile of summed slices transformed into polar coordinates and radial noise-to-  
37 signal ( $R_{\text{NSR}}$ ) also calculated.

38 **Results:** Artifacts were considered sufficiently severe to require camera retuning in 10 rods  
39 sections, 17 sphere sections, and 16 uniform sections. In uniform sections, there was “good  
40 agreement” for inter-observer and intra-rater assessments ( $\kappa = 0.66$ , Fisher exact  $p < 0.0001$  and  
41  $\kappa = 0.61$ , Fisher exact  $p = 0.001$ , respectively). While 3 radiomics image analysis features agreed  
42 significantly ( $p = 0.001$ ) with visual detection of significant artifacts in uniform sections, the  
43 parameters most strongly associated with severe artifacts were  $R_{\text{Contrast}} > 4.75\%$  and  $R_{\text{NSR}} >$   
44  $2.7\%$ , for which ROC AUC accuracy =  $88\% \pm 5\%$ , sensitivity =  $83\%$ , specificity =  $83\%$ ,  $p <$   
45  $0.0001$ . Accuracy was  $76\%-78\%$  for the 3 radiomics features, with significantly lower specificity

46 (48%-61%,  $p < 0.05$ ) than  $R_{\text{Contrast}}$  and  $R_{\text{NSR}}$ . Increasing magnitude of  $R_{\text{Contrast}}$  and  $R_{\text{NSR}}$   
47 correlated significantly with increasingly severe artifact scores ( $\rho = 0.71-0.72$ ,  $p < 0.0001$ ).

48 **Conclusion:** There was good agreement among physicists as to the presence of circular ring  
49 artifacts in uniform sections of SPECT quality assurance scans, with artifacts accurately detected  
50 by radial contrast and noise-to-signal ratio measurements.

51

52 **Keyword:** Image artifacts; SPECT; gamma camera; quality assurance; image texture analysis;  
53 inter-observer agreement; phantom

54

55

56

57

58

59

60

61

62

63

64

65

66

67

68

69

70

71 **Abbreviations**

72 AUC area under curve

73 GLRLM Gray Level Run Length Matrix

74 GLCM Gray Level Co-occurrence Matrix

75 LASSO logistic least absolute shrinkage and selection operator

76 QA quality assurance

77  $R_{\text{Contrast}}$  radial contrast

78  $R_{\text{NSR}}$  radial noise-to-signal ratio

79 ROC Receiver operating characteristics

80 SPECT single photon emission computed tomography

81 VOI volume of interest

82

83

84

85

86

87

88

89

90 **Declarations**

91 Ethics approval and consent to participate: Not applicable.

92 Consent for publication: Not applicable.

93 Availability of data and material: The data sets used for the current study are available for the  
94 corresponding author on reasonable request.

95 Competing interests: None of the authors have any competing interests to report relevant to this  
96 investigation.

97 Funding: Not applicable.

98 Author's contributions: EM, AS, GB and KJN created and executed image analysis algorithms  
99 and performed statistical analysis of results. FPD and KJN performed visual readings. All  
100 authors contributed meaningfully to study design and to writing the manuscript. All authors have  
101 read and approved the final manuscript.

102 Acknowledgements: The authors Fitzgerald Leveque, May Liu, and Christopher Caravella for  
103 their invaluable assistance in collecting and processing the phantom data analyzed in this  
104 investigation.

105

106

107

108

109

**110 Background**

111 Numerous factors influence the accuracy of single photon emission computed tomography  
112 (SPECT) reconstructions. Besides patient motion issues (e.g., tidal breathing motion and cardiac  
113 contraction), complications include photon attenuation and scatter, variability of point spread  
114 function with distance from detectors, imperfect detector efficiency, and low counting statistics  
115 [1]. Over the years considerable efforts have been applied to address these issues through the  
116 application of physics to optimize the reconstruction process and electric and mechanical  
117 engineering to improve SPECT system designs. Yet, data acquisition problems can occur  
118 abruptly, or as a result of gradual electrical or mechanical deterioration of SPECT system  
119 components, which can compromise the reliability of acquired data for clinical use [2].

120 For this reason, adherence to regularly scheduled standardized quality assurance (QA)  
121 procedures is important in order to identify problems with data acquisitions before they can  
122 potentially result in inaccurate clinical readings. These QA procedures include quarterly SPECT  
123 phantom reconstructions, which typically are assessed visually for contrast, spatial resolution,  
124 and uniformity of response [3].

125 In the course of evaluating SPECT phantom transaxial reconstructions to assess integrity of  
126 acquired image data, the appearance of conspicuous concentric ring artifacts often will trigger  
127 retuning conventional rotating SPECT detectors, as those artifacts may be due to inadequate  
128 uniformity corrections [4]. One cause of this can be insufficient count density of the correcting  
129 flood field, the details of which have been studied extensively [5]. It has been shown that choice  
130 of tomographic reconstruction filters affects the appearance of artifacts, and that perceptibility of  
131 artifacts depends on their size and position within the object [6].

132        However, these evaluations are visual, not quantitative, and criteria have not been established  
133 for deciding what constitutes an artifact that is sufficiently severe to necessitate detector  
134 retuning. When medical imaging physicists convene to read QA studies from other institutions,  
135 they begin with training data sets to calibrate their readings and to minimize perceptual  
136 differences among them, and they read by consensus. However, different viewing environments,  
137 such as brightness, contrast and screen resolution of display monitors, can influence visual  
138 impressions. Also, technologists who acquire QA data may not be as well versed in these types  
139 of artifacts may be uncertain as to just how serious a potential artifact could be.

140        To address these issues, algorithms have been devised to analyze standardized SPECT QA  
141 phantoms, both to report the results of spatial resolution and contrast results [7,8], and to report  
142 the possibility of artifacts [9]. Some of these algorithms are based on texture analysis [7,9].  
143 Approaches to quantifying the severity of circular ring artifacts have included annular sampling  
144 techniques [10], or by the use of the Hough transform, the efficacy of which to detect artifacts  
145 has been assessed using computerized Monte Carlo phantom simulations [11].

146        Our study was conducted with two goals: (1) to determine how well medical physicists who  
147 are experienced in making these evaluations agree with one another about the presence and  
148 severity of concentric ring artifacts, and (2) to establish if there are any quantitative image  
149 metrics that correspond to visual judgments as to the significance of these artifacts. Having a  
150 quantitative parameter could help less experienced observers decide whether to escalate a  
151 remediation process when confronted with an imperfect-appearing phantom reconstruction  
152 result.

153

154 **Methods**155 **Phantom data acquisitions**

156 All test data were acquired at one institution as part of routine quarterly QA using standardized  
157 SPECT phantoms. The standardized phantom included a cylindrical water bath, with Plexiglas®  
158 inserts of 6 rod sizes, solid spheres of 6 sizes, and a uniform volume devoid of rods or spheres  
159 [12,13]. Phantoms were loaded with 666 MBq-740 MBq  $^{99m}\text{Tc}$ . Each 360° set of projection data  
160 were comprised of 128 projections acquired as 128x128 matrices for 32-36 x 10<sup>6</sup> counts.  
161 Magnification factors of 1.00 to 1.85 were used (pixel size = 3.3±0.7 mm; range = 2.1 to 5.2  
162 mm).

163 All data sets were reconstructed by filtered backprojection with a Hanning post-filter (cutoff  
164 = 1.0 cm<sup>-1</sup>), and corrected by a simple Chang attenuation correction with attenuation coefficient  
165 of  $\mu = 0.11 \text{ cm}^{-1}$ , or as adjusted as necessary in order to obtain the most uniform appearance of  
166 transaxial reconstructions of the volume of the phantom containing no rods or spheres.

167 From a collection of over 200 acquisitions performed for 12 different SPECT systems  
168 between 2016-2020, 40 studies were identified by a medical physicist at the time of data  
169 acquisition as either having evidence of concentric ring artifacts, or else were acquired to  
170 determine whether concentric ring artifacts were resolved following camera retuning performed  
171 within 1 week after obtaining an unacceptably non-uniform-appearing result. A 41<sup>st</sup> phantom  
172 study also was considered as a reference standard for which observers agreed that no artifacts  
173 were visible. At the time of these QA phantom data acquisitions, camera ages ranged from 20  
174 years old to 1 month old. Ten systems were dual-detector systems and 2 were single-detector  
175 systems.

176

177 **Visual phantom readings**

178 Automated algorithms written in IDL 8.4 were run on all transaxial phantom data that generated  
179 a series of output jpg image files [7]. Because some accrediting agencies request displays of all  
180 transaxial slices [14], multiple jpg files were generated per phantom for all transaxial slices  
181 spanning the entire height of the phantom, and a separate composite jpg file was generated  
182 showing a composite of 3 images (Fig. 1): a summed rod section, the single section through  
183 spheres with the highest contrast, and the section of uniform activity concentration. The jpg files  
184 were reviewed and scored by observers as to the presence and severity of concentric ring  
185 artifacts.

186 To gauge inter-rater agreement of the presence and severity of concentric artifact, two  
187 experienced medical physicists (KJN and FPD) independently scored severity of artifacts on a 5-  
188 point scale (0 = “no artifact” ; 1 = “probably insignificant artifact”; 2 = “equivocal”; 3 =  
189 “probably significant artifact”; 4= “severe artifact definitely requiring camera retuning”),  
190 separately for rod sections (Fig. 2), sphere sections (Fig. 3) and uniform phantom sections (Fig.  
191 4). Single, separate, scores were generated by both readers for the perception of artifacts as seen  
192 on the composite summary images (Fig. 1).

193 One physicist (KJN) re-scored the images blinded to his initial scores, to enable assessment  
194 of intra-rater reproducibility. Readers had no knowledge of any texture analysis metrics values,  
195 as these were not computed until after all readings were completed for both physicists.  
196 Following the assignment of all scores to all phantoms by both physicists, readings also were  
197 dichotomized so that any reading  $> 2$  was assigned 1 and were otherwise 0.

198

199

**200 Radiomic features extraction**

201 Texture analysis was performed by two medical physicists (EM and AS) on DICOM files of the  
202 transaxial reconstructions that were transmitted electronically to them in Italy. SPECT datasets  
203 were imported into MIMvista software (MIM Software, Inc. Cleveland, OH). Using the open  
204 source CGITA software [15], they manually positioned cylindrical volumes of interest (VOI) to  
205 sample counts of the transaxial phantom sections of uniform activity concentration, in the region  
206 of the phantom containing no rods or spheres (Fig 1). When visible, VOIs were centered on a  
207 perceived center of a concentric ring artifact, individually for each phantom data set; otherwise,  
208 they were centered on the middle of a transaxial section. The CGITA algorithms enabled the  
209 computation of 72 image texture features (Table 1), which included image texture metrics based  
210 on Gray Level Co-occurrence Matrix (GLCM) computations [16,17], including image energy  
211 and entropy measures that have been useful in tumor characterization studies [18]. A GLCM is  
212 a matrix that defines the distribution in an image of co-occurring count levels per voxel at a  
213 given voxel offset. Among other additional metrics, the CGITA algorithms computed 11  
214 different types of features of voxel-alignment metrics (Table 2) and 11 intensity size zone  
215 metrics (Table 3) [9]. Voxel-alignment features are useful in tumor characterization studies [19],  
216 and are based on the construction of a Gray Level Run Length Matrix (GLRLM) for statistical  
217 image texture characterization that consists of counting the number of voxels having the same  
218 count level in a given direction [20]. CGITA analysis parameters were selected to establish the  
219 size of GLCM voxel neighborhoods and the extent and orientation angles of the GLRLMs as  
220 follows: (1) local normalization was performed with minimum and maximum count values  
221 calculated inside the VOI; (2) the distance between pairs of voxels was set to 1 voxel (thereby

222 specifying that computations were to be performed for directly neighboring voxels); (3) the  
 223 angular orientation was averaged; and (4) full range of scaled count values was set to 64 bins.

224

### 225 **Radial features extraction**

226 Taking into account the radial nature of the concentric ring artifacts, two more image features  
 227 were computed: radial contrast ( $R_{\text{Contrast}}$ ) and radial noise-to-signal ratio ( $R_{\text{NSR}}$ ) [21].  $R_{\text{Contrast}}$  and  
 228  $R_{\text{NSR}}$  were computed from the 41 SPECT phantom images  $f(x,y,z)$  using the “ImageJ” software  
 229 [22]. The workflow adopted to obtain radial features is illustrated in Fig. 5. Depending on voxel  
 230 dimension, 8-10 tomographic homogeneous phantom transaxial sections ( $z_H$ ) were summed  
 231 together:

$$232 \quad f_z(x,y) = \sum_{z_H} f(x,y,z)$$

233 Summed images  $f_z(x,y)$  were transformed to polar coordinates  $f_z(\rho,\phi)$ , where:

$$234 \quad \rho = \sqrt{x^2 + y^2}$$

235 and:

$$236 \quad \phi = \tan^{-1}\left(\frac{y}{x}\right)$$

237 To obtain the radial profile  $f_{z\phi}(\rho)$ , the images  $f_z(\rho,\phi)$  were summed over the  $\phi$  angular variable:

$$238 \quad f_{z\phi}(\rho) = \sum_{\phi} f_z(\rho,\phi)$$

239 This profile was used to compute  $R_{\text{Contrast}}$  and  $R_{\text{NSR}}$ :

$$240 \quad R_{\text{Contrast}} = \frac{\text{Max}\left(f_{z\phi}(\rho)\right) - \text{Min}\left(f_{z\phi}(\rho)\right)}{\text{Max}\left(f_{z\phi}(\rho)\right) + \text{Min}\left(f_{z\phi}(\rho)\right)}$$

241

$$R_{NSR} = \frac{\sigma(f_{z\phi}(\rho))}{\mu(f_{z\phi}(\rho))};$$

with  $\mu$  and  $\sigma$  being the mean value and the standard deviation of the radial profile  $f_{z\phi}(\rho)$ , respectively.

## Features analysis

In addition to metrics being computed individually for each phantom data set, texture features were also compared to a reference standard image data set. To do so, all metrics were computed for a reference phantom that had been identified as being *a priori* uniform, i.e., graded as “0” independently by both physicists for all rod, sphere, and uniform sections. Relative differences then were computed between the reference standard (i.e. the artifact-free phantom data set) and each of the 40 test data sets.

## Statistical analysis

Statistical analysis was performed using MedCalc software [23]. Values are reported as means  $\pm$  one standard deviation. Whether or not continuous variables were normally distributed was determined by the Shapiro-Wilk test. Trends of variables with increasing severity of concentric ring artifacts were quantified by  $\chi^2$  rank correlation with Spearman’s  $\rho$ . Significance of mean differences was assessed by ANOVA, which included computation of F-ratios (measuring variations among groups for which a value of 1.0 results for the null-hypothesis). The most discriminant of the 72 radiomic features were identified by means of the logistic least absolute shrinkage and selection operator (LASSO) approach [24]. LASSO is a regression analysis method that both selects discriminant variables and performs regularization in order to

264 enhance the accuracy of model predictions and interpretability of the statistical models that are  
265 produced. The LASSO operator for our application used the dichotomous classification of the 40  
266 phantom images, and a shrinkage penalty parameter ( $\lambda$ ) [25]. Inter-rater agreement and intra-  
267 rater reproducibility of visual scores was determined by the kappa statistic [26], with significance  
268 determined by Fisher's exact test. Dichotomized visual scores were formed as  $> 2$  of the mean of  
269 the 2 readers' severity grades of the uniform phantom sections. Logistic regression determined  
270 which of the radiomic and radial features were able to discriminate positive from negative  
271 dichotomous visual scores.

272 Receiver operating characteristics (ROC) analysis determined the accuracy of each  
273 parameter to agree with dichotomous visual scores, producing the threshold for discrimination  
274 (Youden index), sensitivity and specificity for each parameter. Comparison of proportions was  
275 performed by the  $\chi^2$  test. Significance of differences among ROC AUC results was assessed [27].  
276 A probability ( $p$ )  $< 0.05$  was considered to be statistically significant.

277

## 278 **Results**

### 279 **Visual scores**

280 Count levels were similar among phantom acquisitions, with  $(34 \pm 1) \times 10^6$  counts, and were not  
281 normally distributed (kurtosis = 6.0,  $p = 0.0004$ ), as they were narrowly clustered near  $34 \times 10^6$   
282 counts, so that differences among phantoms of counts or of noise-to-signal ratios were unlikely  
283 to account for perceived artifacts.

284 Among the 40 phantoms, artifacts were considered sufficiently severe to require camera  
285 retuning based on the visual scores for 10 rods sections, 16 sphere sections, and 21 uniform  
286 sections for the first reader and for 10 rods sections, 17 sphere sections, and 16 uniform sections

287 for the second reader. Thus, the percent of phantoms in which artifacts were perceived were  
288 smaller in rod than in uniform sections (24% versus 39%-51%,  $p = 0.01$ ). From the displays of  
289 all individual transaxial sections, inter-observer comparisons demonstrated “fair agreement” as  
290 to artifacts for rods ( $\kappa = 0.33$ ,  $p = 0.09$ ), “good agreement” for spheres ( $\kappa = 0.64$ ,  $p < 0.0001$ ),  
291 and “good agreement” for uniform sections ( $\kappa = 0.65$ ,  $p < 0.0001$ ). Correlation between readers  
292 was stronger for sphere sections and uniform sections than for rod sections ( $p = 0.004$ ).

293 Intra-observer reproducibility demonstrated “moderate agreement” for rods ( $\kappa = 0.58$ ,  $p =$   
294  $0.0001$ ), “good agreement” for spheres ( $\kappa = 0.70$ ,  $p < 0.0001$ ), and “good agreement” for uniform  
295 sections ( $\kappa = 0.70$ ,  $p < 0.0001$ ).

296 For the perception of artifacts seen on composite images, there was “moderate agreement”  
297 for inter-observer comparisons ( $\kappa = 0.47$ , Fisher exact  $p = 0.01$ ), and “good agreement” for intra-  
298 observer reproducibility ( $\kappa = 0.68$ ,  $p < 0.0001$ ).

299

### 300 **Correlation with increasing artifact severity**

301 The mean of both observers’ scores of uniform phantom sections were used to establish 5-level  
302 consensus readings of artifact severity, rounded up for non-integer mean scores. Of all image  
303 metrics tested,  $R_{\text{Contrast}}$  had the highest Spearman coefficient with increasing magnitude of  
304  $R_{\text{Contrast}}$  correlated to increasing severity of artifacts ( $\rho = 0.72$ ,  $p < 0.0001$ ) (Fig. 6).  $R_{\text{NSR}}$  also had  
305 significant correlation with increasing artifact severity ( $\rho = 0.71$ ,  $p < 0.0001$ ).

306

### 307 **Image analysis metrics**

308 Both radiomic and radial features were normally distributed ( $p\text{-value} > 0.05$ ). For all of them the  
309 absolute value of the difference between the reference standard and the 40 test data sets was in

310 stronger agreement with dichotomous visual scores than the original metrics themselves.  
311 Consequently, the results reported in Tables 4–6 are for absolute values of differences from the  
312 reference standard phantom.

313 Among all tested parameters,  $R_{\text{NSR}} > 2.7\%$  and  $R_{\text{Contrast}} > 4.75\%$  had the highest accuracy  
314 (ROC AUC =  $88 \pm 5\%$ ,  $p < 0.0001$ ), and similar sensitivity and specificity of 83% (Table 4), with  
315 dichotomous accuracy of 83% for discriminating between positive and negative mean  
316 dichotomous scores. In addition, LASSO indicated that 3 radiomic features (i.e. 1 Intensity-size-  
317 zone feature and 2 voxel-alignment features) significantly predicted dichotomous visual artifact  
318 scores, with a shrinkage penalty parameter  $\lambda = 0.11$ , corresponding to the cross-validation error  
319 between the model and the dichotomized visual scores [25]. For these 3 radiomic features, ROC  
320 analysis indicated  $p < 0.05$  (Fig. 7). For a ROC result being significantly different from chance  
321 (ROC AUC = 50%), the metric's AUC would have to have been  $> 70\%$ . While ROC AUC and  
322 sensitivity values were not statistically different among the parameters reported in Table 4,  
323 specificity was significantly lower for the 3 radiomic features than for  $R_{\text{NSR}}$  and  $R_{\text{Contrast}}$ .  
324 Moreover,  $R_{\text{NSR}}$  and  $R_{\text{Contrast}}$  had the highest ROC AUC values (Table 4).

325 The metric with the strongest association (highest logistic regression  $\chi^2$  value) with visual  
326 scores of artifacts seen in uniform sections was  $R_{\text{NSR}}$ , for which the logistic regression model  
327 correctly predicted 83% of cases, with similar results for  $R_{\text{Contrast}}$  (Table 5).

328  $R_{\text{NSR}}$  and  $R_{\text{Contrast}}$  had significantly higher values for phantoms with severe uniformity  
329 concentric ring artifacts than those without artifacts (Table 6, Fig. 8). Consistent with ROC and  
330 logistic regression analyses results, the F-ratios were higher, and the probability values lower, for  
331  $R_{\text{NSR}}$  and  $R_{\text{Contrast}}$  than for the radiomic features (Table 6).

332

333 **Discussion**

334 As with any type of device used to collect data that contributes to forming a clinical diagnosis, it  
335 is important to assess the performance characteristics of rotating SPECT systems. SPECT  
336 equipment testing includes the acquisition and analysis of three-dimensional phantoms to gauge  
337 system tomographic spatial resolution, image contrast, and uniformity of response, and it is  
338 recommended to be performed as part of acceptance testing [28], and thereafter on a quarterly  
339 basis [3]. Failure to recognize concentric ring artifacts can contribute to misinterpretation of  
340 nuclear cardiology studies [29], and non-cardiac nuclear medicine studies [5].

341 Establishing whether a particular image feature is visible is non-trivial. One of the goals of  
342 our investigation was to document the extent to which two independent medical imaging  
343 physicists agree on the appearance and severity of artifacts. Thresholds of visible detection in a  
344  $34 \times 10^6$  count phantom study and in a clinical study might be different. Experience of the  
345 observer is a factor, and evaluations even by experienced physicists might differ. Visibility  
346 indexes can help in this regard, such as those dependent on feature contrast and feature  
347 dimensions [30], but ultimately such indices require calibration with scores assigned by human  
348 observers [31].

349 We found that readers agreed more successfully when observing all of the transaxial sections  
350 than when viewing just the summary images, and that our readers had “good agreement” in  
351 detecting artifacts in uniform and sphere sections but only “fair agreement” about rods sections.  
352  $\kappa$  values of reader agreement and reading reproducibility were higher for uniform sections and  
353 for sphere sections than for rod sections. That may be due to the fact that there is less distraction

354 in uniform sections in perceiving ring artifact patterns compared to rod sectors, for which such  
355 patterns are superimposed on count variations within and between rod images of different sizes.

356 Visibility of a concentric ring artifact is one thing, but deciding that an artifact is sufficiently  
357 severe to cause potential clinical reading problems is another matter, because that requires  
358 removing a camera from clinical service long enough to do a thorough re-tuning of the detectors.  
359 The process of camera re-tuning may be a fairly unobtrusive process, such as routine detector  
360 tuning or flood calibration by a technologist of  $100-120 \times 10^6$  counts, or can be more time-  
361 consuming, involving a full calibration of energy, linearity, and uniformity corrections by a field  
362 service engineer. Consequently, having a quantitative basis for triggering the decision to retune  
363 detectors would be useful information for technologists performing quarterly SPECT quality  
364 assurance scans.

365 While the results of quality assurance phantom scans usually are assessed visually, recent  
366 progress has been reported in quantifying scanner performance automatically [7,8,9,11]. Among  
367 the parameters that are evaluated in SPECT phantom scans, non-uniformity is perhaps the most  
368 challenging. O'Connor et al [6] found that choice of reconstruction filters has a pronounced  
369 effect on the manifestation of planar flood field non-uniformity on the formation of tomographic  
370 concentric ring artifacts, and that the ability of observers to perceive such artifacts depends on  
371 the size and location of these artifacts; smaller artifacts closer to the center-of-rotation were less  
372 perceptible than larger ones further from the center [6]. That is consistent with a reader tending  
373 to give more credence to a pattern perceived to be repeated over more pixels than over fewer  
374 pixels, as there would be a tendency to give more weight given to a similar pattern confirmed in  
375 multiple locations, even when these deviations from neighboring counts are small. Concentric  
376 ring artifacts constitute a type of pattern, and as such, the bigger the radius, the more likely a

377 human observer will detect it in the first place, even if subtle. The larger the size of the pattern,  
378 the more likely a person is to perceive it, as the suggestion of a pattern in one area of a field can  
379 then be reinforced in neighboring areas, and the more areas that confirm the same impression,  
380 and that conform to a continuation of a pattern in neighboring areas, the higher the confidence of  
381 the reader that the pattern is genuine.

382 An advantage of the polar coordinate transformations used in our study, and in the  
383 application of the Hough transform implemented by Hirtl et al [11], is the fact that non-  
384 uniformities tend to be concentric rings, so that ring artifacts of any size become mapped into  
385 straight lines across the entire extent of the image matrix in polar coordinates. Polar coordinate  
386 transformations have the effect of bringing concentric artifacts that are near to the center-of-  
387 rotation up to an equal footing with those at a greater radius. This process relies on knowing the  
388 center of the reconstruction matrix (i.e. if an offset was applied during image reconstruction). A  
389 possible area of further study could be the implementation of a search algorithm that maximizes  
390 radial contrast ( $R_{\text{Contrast}}$ ) and/or radial noise-to-signal ( $R_{\text{NSR}}$ ), in an attempt to locate this matrix  
391 center.

392 Using ROC analysis, we independently found that a radial contrast value  $> 4.75\%$  confirmed  
393 findings of other investigations that  $> 4\%$  of the amplitude of Hough-transformed transaxial  
394 images above background counts agreed optimally with medical physicists as to the seriousness  
395 of concentric ring artifacts [11]. The fact that the  $R_{\text{Contrast}}$  and  $R_{\text{NSR}}$  parameters were more  
396 effective at agreeing with our medical physicists scores than any of the 72 radiomics features  
397 likely can be explained by the fact that the radial parameters make use of the *a priori* condition  
398 that these artifacts are circular patterns, whereas the radiomics features are more generic and  
399 shape-independent.

400 **Limitations**

401 To maximize spatial resolution, patient data usually are acquired with contouring, not as circular  
402 orbits. Our acquisitions were performed with contouring, and while some acquisitions may have  
403 been of a fixed radius, they were not necessarily so. Considering that the phantom we used was a  
404 simple cylinder, these orbits likely were close to being circular. It has been shown that non-  
405 uniform planar flood field corrections create circular artifacts in SPECT transaxial  
406 reconstructions for orbits of a fixed radius, but more complicated shapes when elliptical orbits  
407 are used [5].

408 Our study was limited to documenting the degree to which medical physicists agree with one  
409 another in their perception of the presence and severity of artifacts, not in whether there were  
410 problems in actual clinical studies. By linking  $R_{NSR}$  and  $R_{Contrast}$  to their visual impressions, we  
411 have provided a means to connect quantified imaging parameters computed from standardized  
412 phantom scans to whether or not an artifact would have been judged to be sufficiently serious to  
413 require service on a SPECT system. This does not guarantee that a SPECT system for which  
414  $R_{Contrast} > 4.75\%$  would invariably result in erroneous diagnoses in all clinical studies  
415 subsequently acquired with this equipment. There are innumerable variables involved in forming  
416 a final clinical diagnosis. Our results just serve as a guideline for acting on the results of phantom  
417 scans.

418 Our analyses of radiomic features and radial parameters were performed only in uniform  
419 sections, not in sphere or rod transaxial sections. The fact that inter-rater and intra-rater  
420 agreement were lowest for rod sections suggest that it is more difficult to decide about the  
421 presence and severity of artifacts in those sections compared to sphere and uniform phantom  
422 sections. Yet, we found the concentric ring artifacts were detected by both readers throughout the

423 individual phantom sections. The fact that some radiomic features agree well with visual  
424 impressions in uniform sections suggests that further work is warranted to enable texture analysis  
425 applications to non-uniform phantom areas containing spheres and rods, where artifacts also  
426 appear.

427 It is possible that re-sampled images or images acquired with different spatial resolutions  
428 may have produced different ROC results from those reported in Table 4 for F1, F2 and F3.  
429 Because of the linear nature by which the radial measures were obtained in our investigation, it is  
430 plausible that  $R_{NSR}$  and  $R_{Contrast}$  should be relatively less sensitive to resampling issues and  
431 relatively less affected by different acquisition parameter choices than the other radiomics  
432 features we computed. The values of textural radiomic features can be affected by acquisition  
433 and reconstruction parameters, as reported for CT and PET imaging [32,33,34], and these  
434 influences require further investigation for SPECT imaging.

## 435 **Conclusion**

436 There is good agreement among observers, with reproducible results, as to the presence of  
437 circular ring artifacts in uniform sections of SPECT quality assurance scans, and some texture  
438 analysis metrics agree well with visual impressions in uniform sections, with highest test  
439 performance characteristics for radial noise-to-signal and radial contrast parameters.

440

441

442

443

444

445

446

447 **Figure Legends**

448 **Fig. 1** The composite summary image for the phantom for which both readers agreed there were  
449 no perceptible artifacts in any phantom section.

450 **Fig. 2** Example of phantom for which both readers scored significant concentric ring artifacts in  
451 sphere sections.

452 **Fig. 3** Example of phantom for which both readers scored significant concentric ring artifacts in  
453 rod sections.

454 **Fig. 4** Example of phantom for which both readers scored significant concentric ring artifacts in  
455 uniform sections.

456 **Fig. 5** Workflow adopted to obtain radial features from SPECT data sets.

457 **Fig. 6** Increasing radial contrast values with increasing values of readers' mean scores for  
458 severity of artifacts.

459 **Fig. 7** ROC curves for identifying phantoms with serious concentric ring artifacts in uniform  
460 phantom sections, for  $R_{NSR}$ ,  $R_{Contrast}$ , F1 (i.e. Voxel-alignment high-intensity short-run emphasis),  
461 F2(i.e. Voxel-alignment low-intensity long-run emphasis) and F3(i.e. Intensity-size-zone high-  
462 intensity short-zone emphasis).

463 **Fig. 8.** Data distributions and box charts, with mean ( $X$ ), percentiles and standard deviations  
464 (SD) for  $R_{NSR}$  (A),  $R_{Contrast}$  (B), F1 (C), F2 (D) and F3 (E) computed for SPECT phantoms with  
465 and without uniform sections concentric ring artifacts.

466

467 **Table 1** CGITA software radiomics classes and features [15].  
 468

Number of computed features	Radiomic class	Feature measure
7	Co-occurrence	second angular moment, contrast, entropy, homogeneity, dissimilarity, inverse difference moment, correlation
11	Voxel Alignment	short-run emphasis, long-run emphasis, intensity variability, run-length variability, run percentage, low-intensity run emphasis, high-intensity run emphasis, low-intensity short-run emphasis, high-intensity short-run emphasis, low-intensity long-run emphasis, high-intensity long-run emphasis
5	Neighborhood Intensity Difference	coarseness, contrast, busyness, complexity, strength
11	Intensity Size Zone	short-zone emphasis, large-zone emphasis, intensity variability, size-zone variability, zone percentage, low-intensity zone emphasis, high-intensity zone emphasis, low-intensity short-zone emphasis, high-intensity short-zone emphasis, low-intensity large-zone emphasis, high-intensity large-zone emphasis
6	Normalized Co-occurrence features	second angular moment, contrast, entropy, homogeneity, dissimilarity, inverse difference moment , correlation
13	SUV Statistics features	minimum SUV, maximum SUV, mean SUV, SUV variance, SUV SD, SUV skewness, SUV kurtosis, SUV skewness (bias corrected), SUV kurtosis (bias corrected), TLG , tumor volume, entropy, $SUV_{peak}$
2	Texture Spectrum	max spectrum, black-white symmetry
4	Texture Feature Coding	coarseness, homogeneity, mean convergence, variance
8	Texture Feature Coding Co-occurrence	second angular moment, contrast, entropy, homogeneity, intensity, inverse difference moment, code entropy, code similarity
5	Neighboring Gray-Level Dependence	small-number emphasis, large-number emphasis, number non-uniformity, second moment, entropy

469

470

471

472

473

474 **Table 2** Voxel alignment extracted features

Voxel alignment feature number	Feature name
1	Short run emphasis
2	Long run emphasis
3	Intensity variability
4	Run-length variability
5	Run percentage
6	Low-intensity run emphasis
7	High-intensity run emphasis
8	Low-intensity short-run emphasis
9	High-intensity short-run emphasis
10	Low-intensity long-run emphasis
11	High-intensity long-run emphasis

475

476

477

478

479

480

481

482

483 **Table 3** Intensity-size-zone extracted features

Intensity-size-zone feature number	Feature name
1	Short-zone emphasis
2	Large-zone emphasis
3	Intensity variability
4	Size-zone variability
5	Zone percentage
6	Low-intensity zone emphasis
7	High-intensity zone emphasis
8	Low-intensity short-zone emphasis
9	High-intensity short-zone emphasis
10	Low-intensity large-zone emphasis
11	High-intensity large-zone emphasis

484

485

486

487

488

489

490

491 **Table 4** ROC analysis test performance characteristics

Parameter	ROC AUC	ROC threshold	ROC p	Sensitivity	Specificity
R <sub>NSR</sub>	88±5%	>2.7%	<0.0001	83%	83%
R <sub>Contrast</sub>	88±5%	>4.7%	<0.0001	83%	83%
F1	78±7%	>0.40	0.0001	89%	61%*
F2	76±8%	>0.077	0.001	100%	48%*
F3	78±7%	>0.34	0.001	94%	52%*

492

493 \* p < 0.05 versus R<sub>NSR</sub>

494 F1 = Voxel-alignment high-intensity short-run emphasis; F2 = Voxel-alignment low-intensity

495 long-run emphasis; F3 = Intensity-size-zone high-intensity short-zone emphasis.

496

497

498

499

500

501

502

503

504

505

506

507

508

509

510

511

512

513

514

515 **Table 5** Logistic regression test performance characteristics  
 516

Parameter	Logistic regression $\chi^2$	Logistic regression p	% cases correctly predicted
R <sub>NSR</sub>	21.8	<0.0001	83%
R <sub>Contrast</sub>	21.2	<0.0001	80%
F1	10.1	0.002	71%
F2	11.0	0.001	71%
F3	9.9	0.002	71%

517

518 F1 = Voxel-alignment high-intensity short-run emphasis; F2 = Voxel-alignment low-intensity

519 long-run emphasis; F3 = Intensity-size-zone high-intensity short-zone emphasis

520

521

522

523

524

525

526

527

528

529

530

531

532

533

534

535

536

537

538

539 **Table 6** ANOVA test performance characteristics  
 540

Parameter	F-ratio	ANOVA p	Negative	Positive
$R_{\text{NSR}}$	26.3	<0.001	2.1±0.7%	3.3±0.7%
$R_{\text{Contrast}}$	24.7	<0.001	3.5±1.3%	5.7±1.5%
F1	10.9	0.002	0.37±0.24	0.58±0.16
F2	10.6	0.002	0.39±0.17	0.54±0.10
F3	10.8	0.002	0.36±0.23	0.57±0.17

541

542 F1 = Voxel-alignment high-intensity short-run emphasis; F2 = Voxel-alignment low-intensity

543 long-run emphasis; F3 = Intensity-size-zone high-intensity short-zone emphasis

544

545

546

547 REFERENCES

- 
1. Rosenthal S, Cullom J, Hawkins W, Moore SC, Tsui BMW, Yester M. Quantitative SPECT imaging: a review and recommendations by the Focus Committee of the Society of Nuclear Medicine Computer and Instrumentation Council. *J Nucl Med.* 1995;36:1489-513.
  2. Celler A, Dixon KL, Chang Z, Blinder S, Powe J, Harrop R. Problems created in attenuation-corrected SPECT images by artifacts in attenuation maps: a simulation study. *J Nucl Med.* 2005;46:335-43.
  3. <https://ACR.org/Clinical-Resources/Accreditation>. Accessed July 31, 2018.
  4. Leong LK, Kruger RL, O'connor MK. A comparison of the uniformity requirements for SPECT image reconstruction using FBP and OSEM techniques. *J Nucl Med Technology.* 2001;9:79-83.
  5. Gullberg GT. An analytic approach to quantify uniformity artifacts for circular and noncircular detector motion in single photon emission computed tomography imaging. *Med Phys.* 1987;14:105-14.
  6. O'Connor MK, Vermeersch C. Critical examination of the uniformity requirements for single-photon emission computed tomography. *Med Phys.* 1991;18:190-7.
  7. Nichols KJ, DiFilippo FP, Palestro CJ. Texture analysis for automated evaluation of Jaszczak phantom SPECT system tests. *Med Phys.* 2019;46(1):262-72.
  8. DiFilippo FP. Assessment of PET and SPECT phantom image quality through automated binary classification of cold rod arrays. *Med Phys.* 2019; 46:3451-61.

9. Sarnelli A, Mezzenga E, Vaghegini A, Piccinini F, Feliciani G, Belli ML, et al. Texture analysis in <sup>177</sup>Lu SPECT phantom images: Statistical assessment of uniformity requirements using texture features. PLoS ONE. 2019;14(7):e0218814.
10. Madsen MT. A method for quantifying SPECT uniformity. Med Phys. 1997;24:1696-700.
11. Hirtl A, Bergmann H, Knausl B, Beyer T, Figl M, Hummel J. Technical Note: Fully-automated analysis of Jaszczak phantom measurements as part of routine SPECT quality control. Med Phys. 2017;44:1638-45.
12. Jaszczak RJ. United States Patent No. 4,499,375; 1985.
13. [http://www.spect.com/pub/Flangeless\\_Jaszczak\\_Phantoms.pdf](http://www.spect.com/pub/Flangeless_Jaszczak_Phantoms.pdf). Accessed July 27, 2018.
14. <https://ACR.org/Clinical-Resources/Accreditation>. Accessed July 7, 2020.
15. Fang YH, Lin CY, Shih MJ, Wang HM, Ho TY, Liao CT, et al. Development and evaluation of an open-source software package "CGITA" for quantifying tumor heterogeneity with molecular images. Biomed Res Int. 2014;2014:248505. doi:10.1155/2014/248505
16. Haralick RM, Shanmugam K, Dinstein I. Textural features for image classification. IEEE Trans Syst Man Cybern. 1973;SMC-3:610-21.
17. Cunliffe AR, Al-Hallaq HA, Labby ZE, Pelizzari CA, Straus C, Sensakovic WF, et al. Lung texture in serial thoracic CT scans: Assessment of change introduced by image registration. Med Phys. 2012;39:4679-90.
18. Dong X, Xing L, Wu P, Fu Z, Wan H, Li D, et al. Three-dimensional positron emission tomography image texture analysis of esophageal squamous cell carcinoma: relationship between tumor <sup>18</sup>F-fluorodeoxyglucose uptake heterogeneity, maximum standardized uptake value, and tumor stage. Nuc Med Commun. 2012;34:40-6.

19. Chu A, Sehgal CM, Greenleaf JF. Use of gray value distribution of run length for texture analysis. *Pattern Recogn Lett.* 1990;11:415-9.
20. Thibault G, Fertil B, Navarro C, Pereira S, Cau P, Levy N, et al. Texture indexes and gray level size zone matrix. Application to cell nuclei classification. 10th International Conference on Pattern Recognition and Information Processing, PRIP 2009, 2009, Minsk, Belarus. pp.140-5. [hal-01499715](#)
21. Michelson, AA. *Studies in Optics.* (University Press, 1927).
22. Rasband WS. ImageJ. U. S. National Institutes of Health, Bethesda, Maryland, USA, <https://imagej.nih.gov/ij/>, 1997-2018.
23. Medcalc Statistical Software version 19.3.1 (MedCalc Software Ltd, Ostend, Belgium; <https://www.medcalc.org>; 2020).
24. Tibshirani R. Regression shrinkage and selection via the Lasso. *J R Statist Soc B.* 1996;58:267–88.
25. Friedman JM, Hastie T, Tibshirani R. Regularization paths for generalized linear models via coordinate descent. *J Stat Softw.* 2010;33;1-22.
26. Landis JR, Koch GG. The measurement of observer agreement for categorical data. *Biometrics.* 1977;33(1):159-74.
27. DeLong ER, DeLong DM, Clarke-Pearson DL. Comparing the areas under two or more correlated receiver operating characteristic curves: a nonparametric approach. *Biometrics.* 1988;44:837-45.

28. Busemann Sokole E, Plachcinska A, Britten A. Acceptance testing for nuclear medicine instrumentation. *Eur J Nucl Med Mol Imaging*. 2010;37:672-81.
29. DePuey EG. How to Detect and Avoid Myocardial Perfusion SPECT Artifacts. *J Nucl Med*. 1994;35:699-702.
30. Rose A. The sensitivity performance of the human eye on an absolute scale. *J Opt Soc Am*. 1948;38:196–208.
31. Burgess AE. The Rose model, revisited. *J Opt Soc Am A*. 1999;16:633-46.
32. Rizzo S, Botta F, Raimondi S, Origgi D, Fanciullo C, Morganti AG, et al. Radiomics: the facts and the challenges of image analysis. *Eur Radiol Exp*. 2018;2(1):1-8.
33. Shafiq-Ul-Hassan M, Zhang GG, Latifi K, Ullah G, Hunt DC, Balagurunathan Y, et al. Intrinsic dependencies of CT radiomic features on voxel size and number of gray levels. *Med Phys*. 2017;44(3):1050-62.
34. Shafiq-ul-Hassan M, Latifi K, Zhang G, Ullah G, Gillies R, Moros E. Voxel size and gray level normalization of CT radiomic features in lung cancer. *Sci Rep*. 2018;8(1):1-9.

# Figures

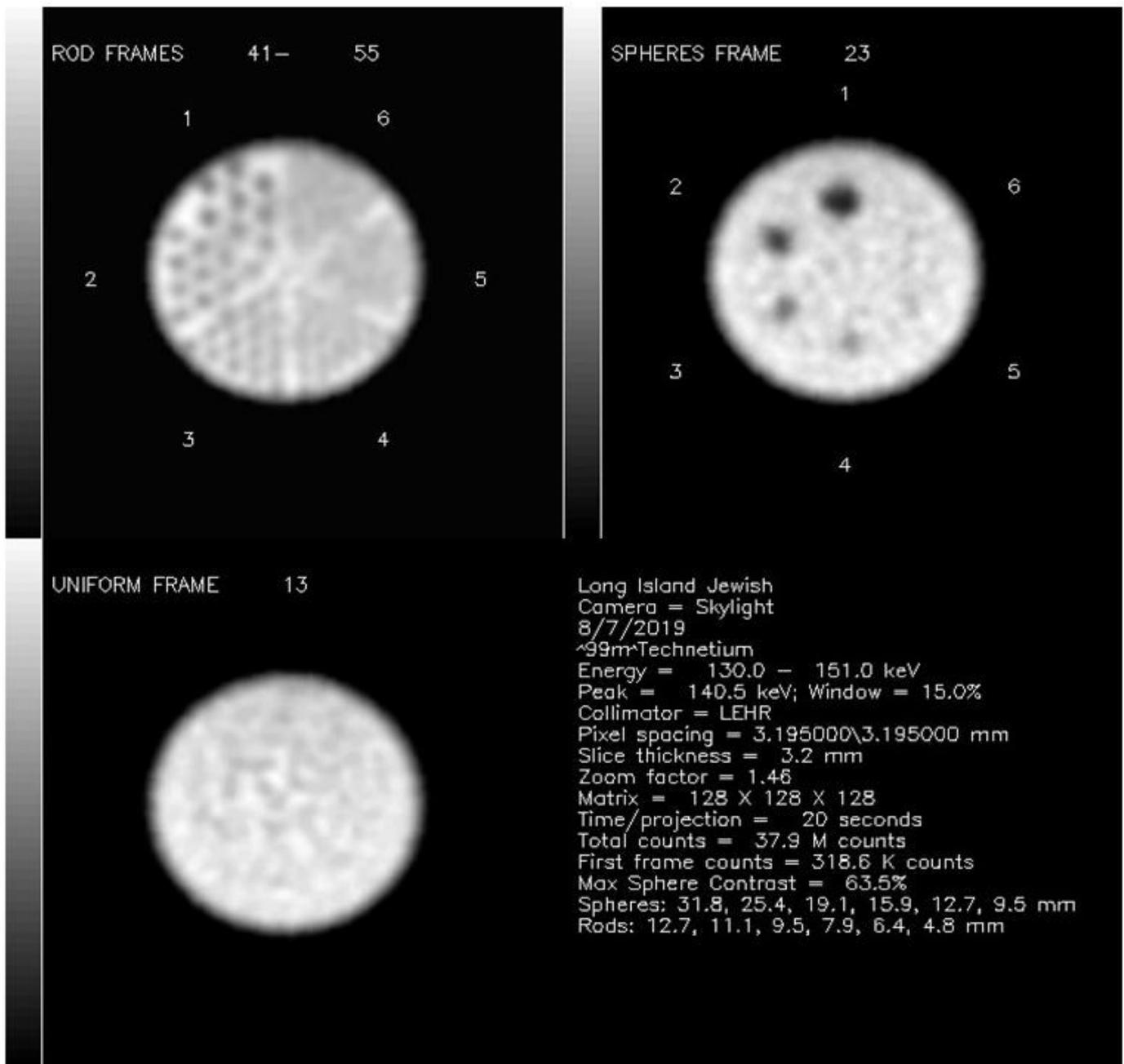
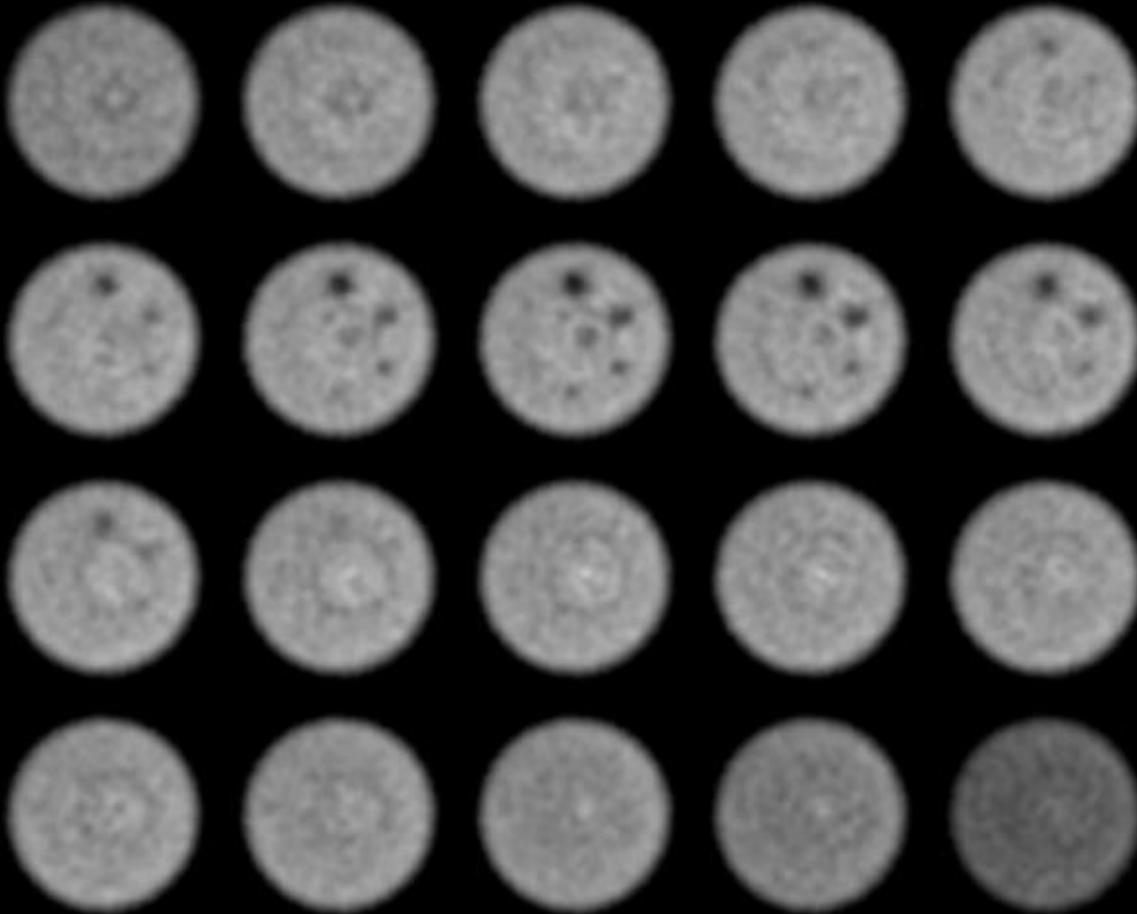


Figure 1

The composite summary image for the phantom for which both readers agreed there were no perceptible artifacts in any phantom section.

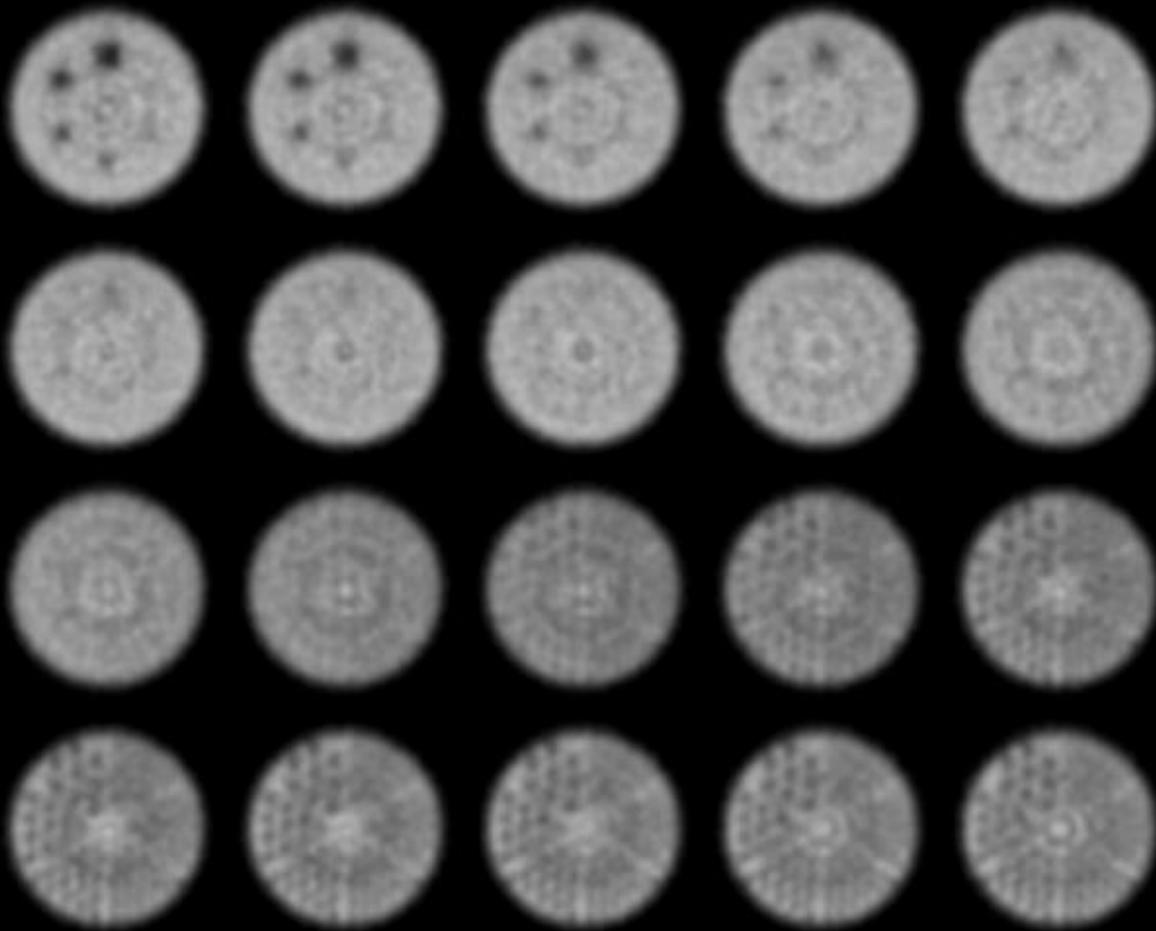
LJ  
Camera = Argus2  
7/1/2016  
<sup>99m</sup>Tc  
Energy = 130.0 - 151.0 keV  
Peak = 140.5 keV; Window = 14.9%  
Collimator = LEHR



**Figure 2**

Example of phantom for which both readers scored significant concentric ring artifacts in sphere sections.

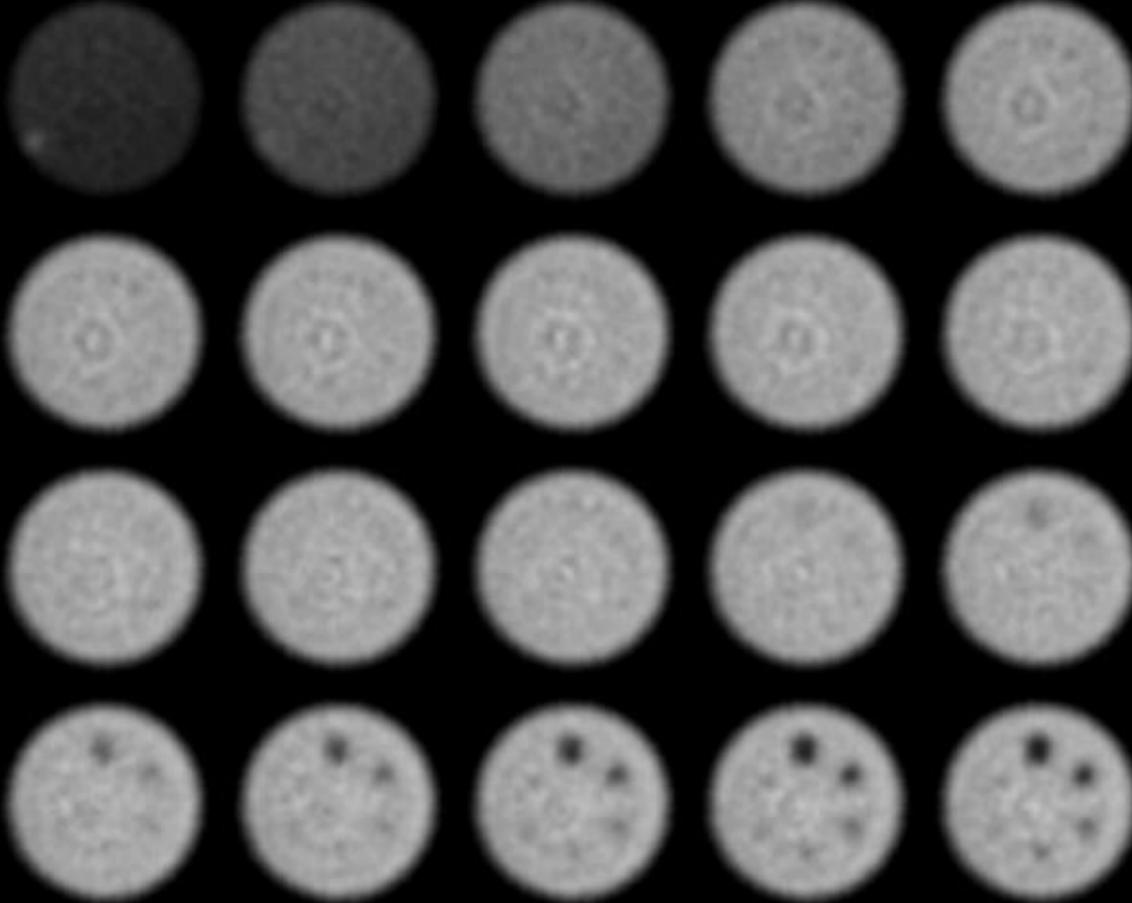
Long Island Jewish Me  
Camera = Argus1  
3/14/2019  
<sup>99m</sup>Tc  
Energy = 129.0 - 150.0 keV  
Peak = 139.5 keV; Window = 15.1%  
Collimator = LEHR



**Figure 3**

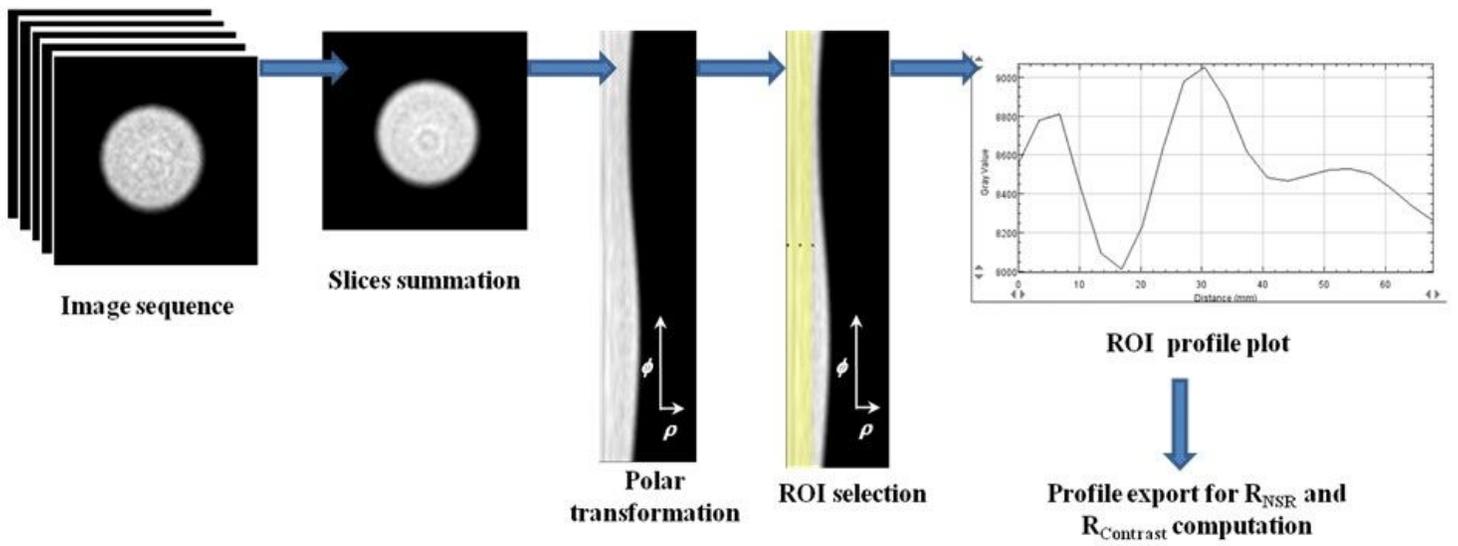
Example of phantom for which both readers scored significant concentric ring artifacts in rod sections.

North Shore/ LJ  
Camera = MILLENNIUM MG  
7/20/2018  
<sup>99m</sup>Tc  
Energy = 129.5 - 150.5 keV  
Peak = 140.0 keV; Window = 15.0%  
Collimator = LEHR



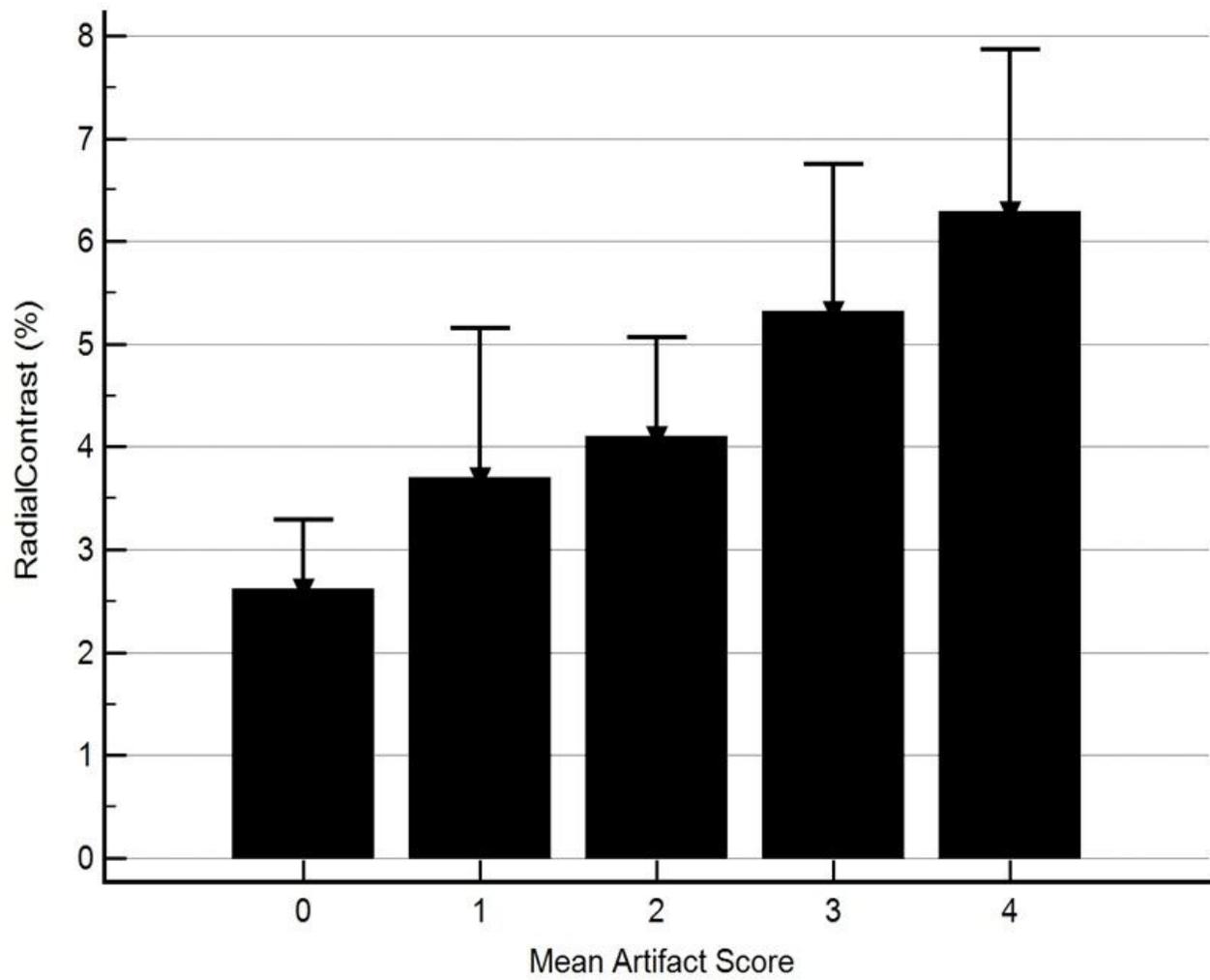
**Figure 4**

Example of phantom for which both readers scored significant concentric ring artifacts in uniform sections.



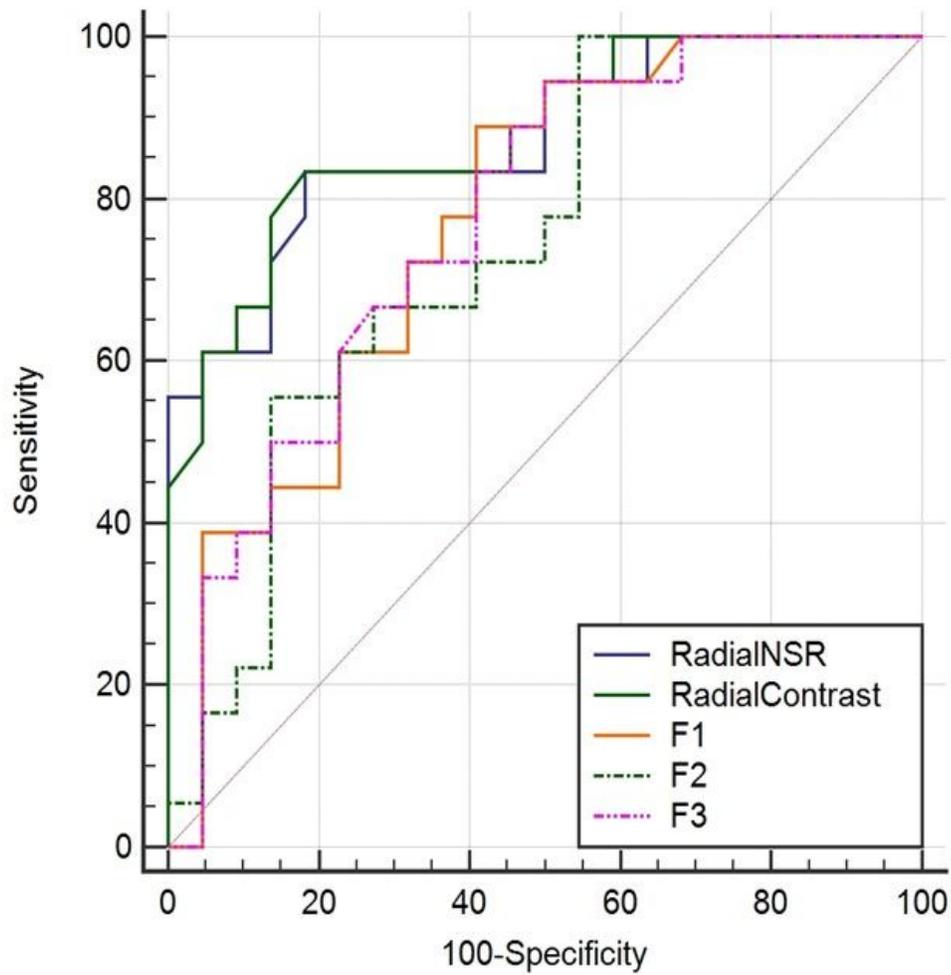
**Figure 5**

Workflow adopted to obtain radial features from SPECT data sets.



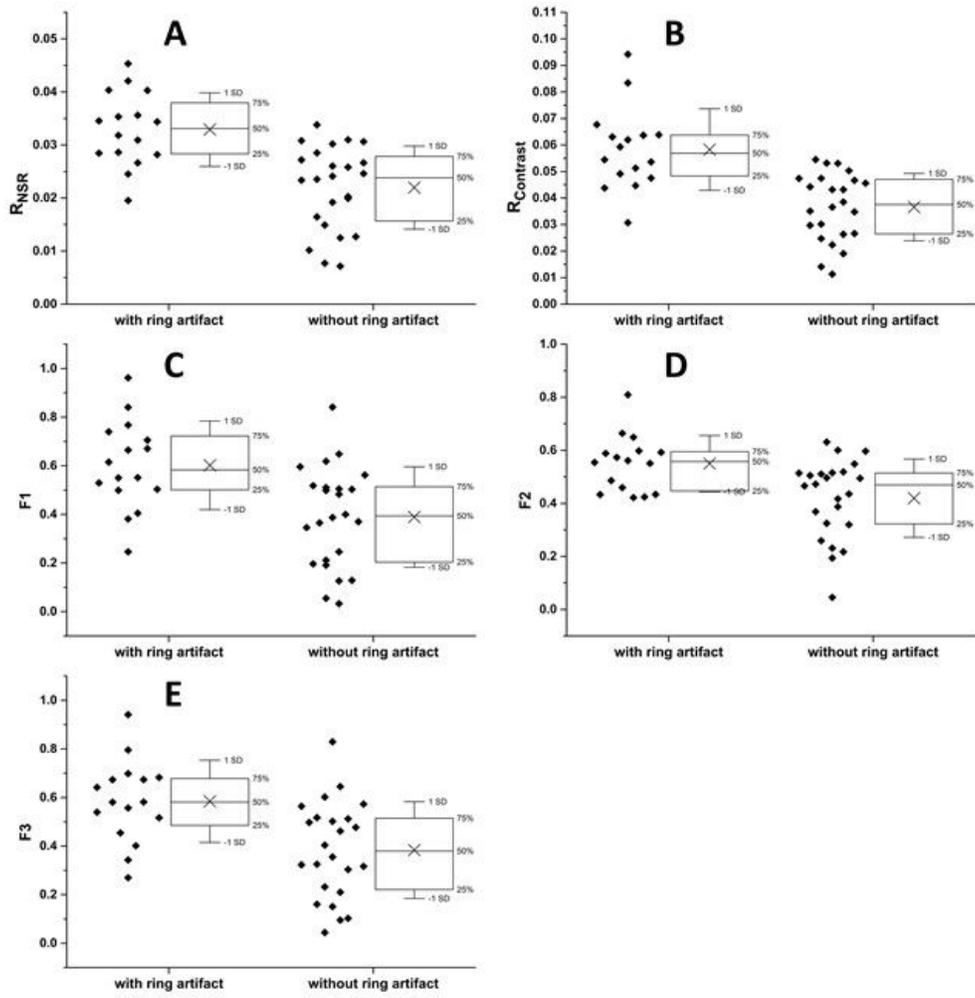
**Figure 6**

Increasing radial contrast values with increasing values of readers' mean scores for severity of artifacts.



**Figure 7**

ROC curves for identifying phantoms with serious concentric ring artifacts in uniform phantom sections, for RNSR, RContrast, F1 (i.e. Voxel-alignment high-intensity short-run emphasis), F2(i.e. Voxel-alignment low-intensity long-run emphasis) and F3(i.e. Intensity-size-zone high-intensity short-zone emphasis).



**Figure 8**

Data distributions and box charts, with mean (X), percentiles and standard deviations (SD) for  $R_{NSR}$  (A),  $R_{Contrast}$  (B),  $F1$  (C),  $F2$  (D) and  $F3$  (E) computed for SPECT phantoms with and without uniform sections concentric ring artifacts.