ANALYSIS OF ASTM X-RAY SHRINKAGE RATING FOR STEEL CASTINGS

Kent Carlson[1], Shouzhu Ou[1],

Richard Hardin[1] and Christoph Beckermann[2]

[1] Research Engineers, Department of Mechanical Engineering, UNIVERSITY OF IOWA, Iowa City, IA
[2] Professor, Department of Mechanical Engineering, UNIVERSITY OF IOWA, Iowa City, IA

ABSTRACT

This paper presents the results of two different studies that examined the ASTM x-ray shrinkage rating system for radiographs of steel castings. The first study evaluated the reliability and repeatability of x-ray shrinkage ratings through a statistical study of x-rays that were each rated seven different times. The second study involved an effort to determine the shrinkage severity level of x-rays through digital analysis of scanned radiographs.

In the first study, a statistical analysis was performed on 128 x-rays that were each given seven ASTM shrinkage x-ray ratings (ratings from five radiographers, two of whom rated all the x-rays twice). The ratings were assigned according to ASTM E186 or E446, depending on which standard was appropriate for each x-ray. It was found that the seven ratings for each x-ray were in unanimous agreement on shrinkage type for 37% of the x-rays, in unanimous agreement on shrinkage level for 17% of the x-rays, and in unanimous agreement for both type and level for 12.5% of the x-rays. All of the x-rays that had unanimous agreement for both type and level were either completely sound or very unsound (Level 5). For each x-ray, the seven shrinkage level ratings were averaged to find a mean level, and a 95% confidence interval was computed to indicate the variance of the ratings. The largest variance was found to occur in x-rays that had a mean x-ray level between 1.5 and 3.5. These x-rays had a 95% confidence interval of about ±2. The average 95% confidence interval for all 128 x-rays was ±1.4. The seven shrinkage levels for each x-ray were then analyzed to determine how many x-rays had shrinkage level ratings both above and below different accept/reject thresholds. This indicates that some of the radiographers would have accepted the castings corresponding to these x-rays, and others would have rejected them. 35% of the x-rays had ratings on both sides of the "Level 1 or better" threshold, 25% had ratings crossing "Level 2 or better", and 21% had ratings crossing "Level 3 or better". Next, the ratings from the two radiographers who rated all 128 x-rays twice were analyzed for consistency. For these radiographers, it was found that their second x-ray rating differed from their first in type at least 19% of the time, in level at least 34% of the time, and in either type or level at least 40% of the time. Both radiographers also reversed accept/reject decisions for 10-15% of the x-rays. This study indicates that x-ray rating is a relatively subjective task, and there is definitely room for improvement in the process.

The second study was undertaken in an effort to quantify the shrinkage x-ray level rating process. ASTM E186 reference radiographs for class CA, CB and CC shrinkage were digitized, and defect measurements were taken. It was found that both the defect area percentage (area of defects/total area) and the defect circumference ratio (circumference of defects/square root of total area) increased with x-ray level for all three classes of shrinkage. Also, for a given x-ray level, these values generally increased with shrinkage class (CA < CB < CC). The increase of defect area and circumference with x-ray level for a given defect class indicates the possibility of determining x-ray level with these measures. Unfortunately, it was found that neither the defect area percentage nor defect circumference ratio could be used to determine x-ray level unless the shrinkage class was already known. Another defect measurement computed for all the reference radiographs was the equivalent defect radius. From the equivalent defect radius results, it was determined that the defect size does not increase appreciably with x-ray level, and that class CA, CB and CC defects all have a size on the order of one to two millimeters.

Next, three trial x-rays were scanned and analyzed to test whether or not the defect area and circumference values could be used to determine x-ray level for radiographs with a known shrinkage class. Several selection areas (defect area plus varying amounts of sound area around the defects) were chosen for each trial x-ray. By comparing both the defect area percentage and defect circumference ratio of each selection area to the corresponding values for the class CB shrinkage reference radiographs, reasonable ASTM shrinkage x-ray levels were determined for each x-ray. Different selection areas did not always indicate the same x-ray level, but considering both defect area and circumference values for all selection areas, it was possible to determine the correct level. The defect area percentage and defect circumference ratio were defined to automatically prorate the area of interest to the area of the reference radiographs. However, the study of different selection areas pointed out the somewhat arbitrary nature of the selection of the "area of interest".

1. Introduction

Radiographic testing is one of the primary measures of casting quality employed in the steel casting industry. Most production castings are either accepted or rejected based on their radiographic testing (RT) levels of various discontinuities, such as gas porosity, inclusions and shrinkage. However, RT levels are determined in a somewhat subjective manner. X-rays of casting sections are assigned a severity level for a particular type of discontinuity through comparison with ASTM standard radiographs for steel castings. These radiographs are included in ASTM standard E446 for castings up to 2 inches thick [1], E186 for castings 2 to 4½ inches thick [2], and E280 for castings 4½ to 12 inches thick [3]. Each of these three standards provides reference radiographs that are grouped by discontinuity category. Category A is assigned to gas porosity, Category B is sand and slag inclusions, and Category C is shrinkage. Shrinkage is further subdivided into types CA (individual veinous strands of shrinkage), CB (grouped or connected veinous strands; almost tree branch-like) and CC (spongy appearance). For sections less than 2 inches thick (E446), there is also a type CD (similar to CB, but more dense and compact). Category A, B and C discontinuities are all divided into five severity levels (1-5, severity increases as the number increases), and there are reference radiographs for each level of each type of discontinuity. In addition to Categories A, B and C, there are also categories for cracks, hot tears, etc. These categories, however, are not divided into severity levels.

To assign a severity level to a particular kind of discontinuity on a production x-ray, it is first necessary to determine what type of discontinuity it is. This is not always trivial, and can be complicated further when more than one type of discontinuity (e.g., shrinkage and sand) is present. Furthermore, if it is determined that the discontinuity is shrinkage, it is also necessary to determine what kind of shrinkage. Once the discontinuity has been categorized, the production x-ray is compared to the reference radiographs for that discontinuity. The highest level of reference radiograph that the production x-ray has a severity better than or equal to is the severity level assigned to that x-ray.

This radiograph comparison process may sound relatively straightforward, but there are several issues that complicate the comparison. First, the discontinuities on a production x-ray may be distributed in a manner that is quite different from the reference radiographs. For example, the reference radiographs for shrinkage have a relatively uniform shrinkage distribution. However, shrinkage on production x-rays is frequently non-uniformly distributed. A good example of this is centerline shrinkage. To determine a severity level in this case, a representative "area of interest" on the production x-ray, containing the centerline shrinkage and an unspecified amount of the surrounding sound region, must be chosen to compare with the standard. The arbitrary choice of the "area of interest," combined with the difficulty in objectively comparing the severity of a uniform pattern to a highly non-uniform one, complicate the determination of severity level. Second, the comparison is supposed to be made using an area of similar size to the reference radiograph. If the area of interest in the production x-ray has a different size, this area is supposed to be prorated to the area of the reference radiograph. This requirement of the radiographer to mentally prorate an area and then compare severity further decreases the objectivity of this determination. Finally, for Category A and B discontinuities, the standards state that the radiographer should not evaluate severity based solely on the size, number or distribution of discontinuities. Rather, the total area of the discontinuities in the production x-ray should be compared to the total area in the reference radiograph. For example, a production x-ray may have larger gas pores than a reference radiograph, but the gas porosity level may still be less severe than the reference if there are fewer pores in the production x-ray. Trying to visually compare total discontinuity areas by considering size, number and distribution is not a simple task.

Because the decision of whether to accept or reject a casting is usually based on these relatively subjective severity levels, it is of interest to quantify the variability involved in assigning RT (or x-ray) levels to a casting x-ray. Quantrell [4] performed a study of the variability of x-ray shrinkage rating using radiographic records from 34 casting defects. Five trained radiographers independently evaluated the x-rays, and the results were compared for consistency in shrinkage type and shrinkage level. With respect to shrinkage type, Quantrell noted that several radiographs had types CA, CB and CC assigned to the same indication. He found that there was more distinction between type CC and the other types, especially at higher severity levels. At lower severity levels, he concluded that type categorizing was very

difficult to do with any degree of reliability. With respect to shrinkage level, 6 of the 34 radiographs were rated with the same level by all five radiographers, and all of those were either Level 1 or Level 5 ratings. Quantrell concluded that Level 3 was the most inaccurately defined level, and that the severity level rating system has a variability of ± 1 level. He also made a noteworthy point, however, in stating that in an accept/reject situation in a foundry, the radiographer only has to compare the production x-ray with one standard radiograph and decide if the defects in the production x-ray are better or worse. For example, if the acceptance criterion is Level 2 or better, it does not matter if the radiographer thinks the production x-ray defect level is Level 3 or Level 4, because both of those would be rejected. This situation should reduce the variability in shrinkage level rating, to some degree.

The next section of this paper describes the results of a study inspired by Quantrell's work. 128 x-rays were rated by five radiographers, two of whom rated all of the x-rays two different times. This study was performed to validate and expand on Quantrell's findings, in order to gain further insight into the reliability and repeatability of x-ray shrinkage ratings. Then in Section 3, a potentially promising method of assessing ASTM shrinkage severity levels through the quantitative analysis of digitized x-rays will be presented.

2. Reliability and Repeatability of X-ray Shrinkage Ratings

The present x-ray shrinkage rating study involved 128 x-rays from the first SFSA / University of Iowa steel plate casting trials [5, 6]. In these casting trials, four foundries cast 3" thick x 6" wide plates, and one foundry cast 1" T x 5.5" W plates. Each foundry cast plates of various lengths that were selected to produce castings ranging from completely sound to RT Level 5 shrinkage. All plates were cast from low alloy steel (1025 steel or similar).

The plate x-rays were originally evaluated for ASTM shrinkage type and severity level by a radiographer at the casting foundry. In the results, this initial x-ray rating is referred to as 'Radiographer 1', even though a different radiographer evaluated the x-rays at each foundry. Then the x-rays from all five foundries were collected and sent to other radiographers for evaluation. Radiographers 2 and 3 each evaluated the x-rays two separate times, and Radiographers 4 and 5 each evaluated the x-rays once. Thus, each x-ray was rated a total of seven times. Radiographers 2 and 3 were Level III Radiographers, and Radiographers 4 and 5 were Level II Radiographers. The x-rays of the 3" T x 6" W plates were evaluated using reference radiographs from ASTM E186, and the x-rays of the 1" T x 5.5" W plates were evaluated with radiographs from E446. Radiographer 2 also evaluated the x-rays for Category A and B discontinuities, and some of the original foundry radiographers provided A and B ratings as well.

The data from the radiographers' evaluations were then collected and analyzed. For each radiograph, the seven x-ray level ratings were averaged to find the mean x-ray level (Xavg). A one-sided student-t 95% confidence interval (tS_x) was also computed for each x-ray, where S_x is the standard deviation of the seven ratings and t = 2.447 is a weighting function accounting for the 95% level of confidence and the small data set size. If the data set size were very large, t would approach 2.0, the value used for 95% confidence intervals for large data sets (two standard deviations). Note that since tS_x is one-sided, the complete 95% confidence interval is twice this value (i.e., the complete confidence interval is Xavg $\pm tS_x$).

In addition to shrinkage severity Levels 1-5, the results presented in this paper add the additional designation of Level 0. Level 0 indicates that the x-ray has absolutely no shrinkage indications (i.e., completely sound). Although this is not an official ASTM classification, making the distinction between completely sound and Level 1 x-rays provides additional information that is useful for the purposes of this study.

2.1 Overall Agreement of X-ray Shrinkage Ratings

The occurrence of unanimous agreement between the seven x-ray ratings can be summarized as follows:

- Unanimous agreement in shrinkage type: 47/128 x-rays (37%)
     14 Level 0 (no type)
     7 CA

20 CB
6 CC

- Unanimous agreement in shrinkage level: 22/128 x-rays (17%)
    14 Level 0
    1 Level 1
    1 Level 2
    1 Level 3
    5 Level 5

- Unanimous agreement in shrinkage type <u>and</u> level: 16/128 x-rays (12.5%)
    14 Level 0
    2 CB5

These results indicate that shrinkage type is more readily identifiable than shrinkage level. Unanimous agreement on x-ray type occurred in nearly two out of every five x-rays, while unanimous agreement on x-ray level occurred in less than one out of every five x-rays. Most of the x-rays with unanimous level ratings were either completely sound (Level 0) or very unsound (Level 5). All of the x-rays with both unanimous type and level ratings were either Level 0 or Level 5. These trends agree with the findings of Quantrell [4], and illustrate the subjective nature of the current x-ray rating system.

Figure 1 gives additional detail regarding the agreement of the x-ray ratings, listing not only the overall results (rightmost category), but also the results for each foundry. Note that Figure 1 shows the percentages of x-rays that do <u>not</u> have unanimous shrinkage type, level and total ratings, which is the opposite of the data presented above. It is evident that there is a significant difference among foundries in the percentage of x-rays without unanimous agreement. The percentage of x-rays without unanimous type agreement ranges from 41% in Foundry K to 94% in Foundry S, while the percentage without unanimous level agreement ranges from 50% in Foundry K to 95% in Foundry E. Recall that Foundry K used E446 as the standard (1" thick plates), while the other four foundries used E186 (3" thick plates). The use of a different set of reference radiographs for the 1" thick plates may contribute to the differences seen in rating agreement.

A listing of the radiographs with the poorest agreement in x-ray level ratings is given in Table 1. It is interesting to note that all but one of these x-rays had significant indications of Category A and/or B discontinuities as well (listed in 'Other Ratings'). It is likely that some radiographers saw discontinuities as either gas or inclusions and thus gave a low shrinkage level, while others determined that the same discontinuities were shrinkage. For x-rays 2 and 6, the foundry radiographers ('Radiographer 1') did not record the shrinkage type, only the level.

Table 1   Worst-case examples of agreement between x-ray ratings

X-ray	Radio-grapher 1	Radiographer 2 1st	2nd	Radiographer 3 1st	2nd	Radio-grapher 4	Radio-grapher 5	Other Ratings
1	0	CB4	CC5	0	CC4	CA1	CA2	B3
2	1	0	0	CC4	CB1	0	0	B2
3	CA2	CB3	CB5	CB1	CA2	CA1	CB3	A3, B3
4	CA3	CB2	CB3	CB1	0	CB4	CB3	none
5	CA2	CB2	CB3	0	CA3	CC4	CB3	B3
6	1	0	0	CC4	CB1	CA1	0	B2
7	0	CB1	CC3	0	CA1	CB4	CB2	B4
8	0	CB4	CB5	CD5	0	CB1	0	A4, B4

2.2 Average 95% Confidence Intervals for X-ray Levels

Figure 2 illustrates the variance of x-ray level ratings for the entire set of 128 x-rays. The x-rays were grouped by average x-ray level, and the one-sided 95% confidence intervals (tS_x values) of all the x-rays in each group were then averaged. The average 95% confidence intervals (CI's) are relatively small for

the x-rays grouped in Levels 0 and 5. This indicates that, on average, a given x-ray in one of these groups has a relatively small variance among its seven x-ray level ratings. Hence, the radiographers were in the best agreement about x-ray level for Level 0 and Level 5 x-rays. The level increases for the middle groups, with the x-rays in Levels 2 and 3 having the largest average 95% CI's. Thus, the x-rays in these groups have, on average, the largest variance in the seven x-ray levels assigned to them. The average 95% CI's of Levels 2 and 3 are 2.1 and 1.9, respectively. This implies that x-rays in these groups have an average variability of about ±2 x-ray levels.

An argument could be made that the eight x-rays listed in Table 1 are outliers, and should not be considered in these results. If these eight x-rays are removed from the data set and the data in Figure 2 are re-plotted, the results do not change drastically. The only levels whose average 95% CI's change are Levels 1 and 2. For Level 1, the average CI drops from 1.7 to 1.5 (for 33 x-rays), and for Level 2, it drops from 2.1 to 1.6 (for 21 x-rays). The overall average 95% CI for all x-rays drops from 1.42 to 1.24. Since the results are similar with or without these x-rays, they are left in for the remainder of the analysis.

The average 95% CI's are displayed again in Figure 3, but this plot groups all of the x-rays from each foundry together. Comparing the average CI's for each foundry, it is evident that there is some difference between foundries. As mentioned with respect to Figure 1, the lower average 95% CI in the x-rays from Foundry K could be related to the use of a different standard (E446) to evaluate those x-rays than those of the other foundries (E186).   Another factor contributing to the difference could be the percentage of x-rays from each foundry that are either completely sound or very unsound. Since the radiographers had better agreement on those x-rays, the average 95% CI for a foundry should decrease with an increase in Level 0 and Level 5 x-rays. This idea is supported by Foundry S (12% of x-rays in Level 0 or Level 5 groupings in Figure 2), Foundry H (30%) and Foundry K (64%). However, Foundries E (28%) and P (55%) seem to have average ranges higher than expected by this logic. It is possible that x-ray quality and casting cleanliness could also play a role in these trends.

2.3   Extent of Disagreement in X-ray Shrinkage Ratings

Further analysis of the variability of the x-ray ratings for each x-ray is presented in Figures 4 and 5, which illustrate the degree of disagreement in level and type ratings, respectively. For each x-ray, the seven RT ratings were compared by level and by type. The most frequently occurring level among the seven level ratings was considered the consensus level, and the most frequently occurring type was considered the consensus type. Figures 4 and 5 group the x-rays by the number of radiographers that disagreed with the consensus level and type, respectively. In Figure 4, the leftmost column shows the 22 x-rays that had a unanimous level rating (zero radiographers disagreed). The next column shows that there were an additional 22 x-rays where only one radiographer disagreed with the consensus. However, in the remaining 84 x-rays, two or more radiographers gave a different level than the rest. Thus, for the 106 x-rays that did not have a unanimous rating, at least two ratings were creating the disparity more than 80% of the time. Figure 5 shows that the radiographers agreed more readily on the shrinkage type, with almost 60% of the x-rays having either a unanimous type or only one dissenting opinion. There were two or more dissenting opinions in 65% of the 81 x-rays that did not have a unanimous type.

2.4   Rating Consistency with Respect to Accept/Reject Threshold Levels

Figure 6 illustrates the variability of x-ray level ratings in terms of common threshold RT levels specified by customers. Since determining whether a casting is better or worse than an accept/reject threshold is the main reason radiographs are produced, this is a critical issue. The three different thresholds specified in Figure 6 are "Level 1 or better", "Level 2 or better" and "Level 3 or better". The columns in this figure represent the percentage of radiographs from each foundry that were assigned x-ray level ratings both above and below the given threshold values. This implies that the castings associated with these x-rays would have been accepted by some radiographers and rejected by others. Note that both Foundries S and E have a threshold category where more than 50% of the castings would be accepted by some radiographers and rejected by others. Considering all threshold categories, Foundry K has the smallest percentage of x-rays with ratings crossing the thresholds. This corresponds to Foundry K having the smallest average 95% CI of x-ray levels (see Figure 3). It seems sensible that as the average 95% CI

increases, the variability of x-ray ratings increases, and hence the number of x-rays with ratings that cross threshold values should also increase. Comparing Figures 3 and 6, this trend is evident. The rightmost column in Figure 6 gives the overall results for all x-rays. At least 20% of the x-rays cross each accept/reject threshold. For "Level 1 or better", 35% of the x-rays cross the threshold.

2.5 Repeatability of X-ray Shrinkage Ratings

Finally, the consistency of the x-ray level ratings by the radiographers who rated all of the x-rays twice was examined. Figure 7 compares the first and second sets of ratings from Radiographer 2. Each x-ray is plotted as a filled circle, with the first rating as the horizontal coordinate and the second rating as the vertical coordinate. The diagonal line passes through the circles representing the x-rays that were given the same rating both times. The numbers next to the circles indicate the number of x-rays that the circles represent. Radiographer 2 was relatively consistent, giving the same level in both ratings 66% of the time. Notice that there are more x-rays above the diagonal line (36) than below (8). Therefore, when Radiographer 2 gave an x-ray a different level in the second rating, the second rating was usually higher than the first. The summary information in the box in the lower right of this plot indicates that Radiographer 2 was consistent in type rating, giving the same type in both ratings 81% of the time. The total rating (type and level) was the same 59% of the time. The dashed boxes enclose x-rays that Radiographer 2 rated once above the "Level 1 or better" threshold, and once below. The dash-dot-dot boxes enclose x-rays rated both above and below the "Level 2 or better" threshold. Radiographer 2 gave x-ray ratings that crossed the "Level 1 or better" threshold for 12% of the x-rays, and crossed the "Level 2 or better" threshold for 10% of the x-rays.

Figure 8 displays the results of the two ratings by Radiographer 3. This radiographer was a little less consistent than Radiographer 2 in shrinkage level (same level 56% of the time), and a considerably less consistent in shrinkage type (same 51% of the time). There is also more scatter in the levels assigned by Radiographer 3. The largest difference in levels between the first and second ratings of Radiographer 3 is five (5 first rating, 0 second), and there are several x-rays with a difference of three levels. By contrast, Radiographer 2 never gave a second rating more than two levels from the first rating. Examining the boxes enclosing x-rays that crossed the two accept/reject thresholds considered, Radiographer 3 crossed the "Level 1 or better" threshold for 15% of the x-rays, and crossed the "Level 2 or better" threshold for 9% of the x-rays.

The information presented in this section indicates that ASTM shrinkage x-ray level rating is definitely not an exact science. For Levels 2 and 3 x-rays, this study found a variability on the order of ±2 levels. The results concerning accept/reject thresholds show that it may not be uncommon for one radiographer to accept a casting while another rejects the same casting. Furthermore, one radiographer looking at one x-ray on two separate occasions may give it an acceptable rating one time and an unacceptable rating the other. In light of this analysis, it seems evident that a more quantitative, less subjective method of determining x-ray levels would be very beneficial to the steel casting industry. The next section describes a study that attempts to develop such a quantitative methodology for the determination of shrinkage x-ray levels.

3. Digitized X-ray Analysis

The objective of the digital x-ray study was to use image analysis of ASTM reference radiographs to establish quantitative relationships between shrinkage defect measurements and ASTM x-ray level. Once established, these relationships could then be used to evaluate x-ray levels from digitized images of casting x-rays in a more objective manner. If successful, this methodology could be used to improve or revise the ASTM standards utilized to determine shrinkage x-ray levels. In the first part of this study, ASTM reference radiographs were digitized, and different defect measurements were evaluated to determine if they correlated to x-ray shrinkage levels. Once correlations were found, the second part of this study consisted of digitizing three production radiographs to determine if it was possible to correctly assess their x-ray level using the correlations developed from the reference radiographs.

3.1 Relating Defect Measurements to X-ray Level for Reference Radiographs

The first step in this study was to convert ASTM standard radiographs to a digital format. ASTM E186 reference radiographs for shrinkage defects were selected for this work. The reference x-rays for shrinkage types CA, CB and CC (Levels 1-5 for each type) were scanned using high optical density scanning equipment at the Iowa State University Center for Nondestructive Evaluation (CNDE). To determine the effects of different scanning exposure times, three of the reference radiographs (CB1, CB2 and CB4) were scanned twice, using different exposure times for each scan to produce different levels of contrast. In addition, the area of significant defects in CB5 was larger than the scanner area, so three different areas of CB5 were scanned to cover the entire image. In the defect measurement plots in this section, if a single value is shown for a reference radiograph that was scanned more than once, this value represents the average value from the two or three scans of that x-ray. The values of the individual scans will also be shown where applicable.

The scanned images of the reference radiographs were processed with image analysis software, and then analyzed to look for correlations between measurable defect attributes and x-ray level. Many different defect measurements were investigated in this work. Three of these measures will be discussed here: (1) defect area percentage (= $A_{defects}/A_{total}$), (2) defect circumference ratio (= $CF_{defects}/sqrt(A_{total})$), and (3) equivalent defect radius (= $2A_{defects}/CF_{defects}$). The scanned images were processed with the software package Scion Image to remove the differences caused by varying x-ray contrast levels and different exposure times during scanning. In order to compute the area, circumference and equivalent radius of defects, it was also necessary to define which parts of the image were defects and which were background. This was done with the software package Transform. The output file from Transform was then read by a FORTRAN program, which calculated the desired defect area percentage, circumference ratio and equivalent radius. An outline of the procedures involved in this study is as follows:

1) Scan x-ray

2) In Scion Image:
   a) Set scale: 5.6 pixels/mm was used in this work
   b) Smooth image: this was done using the option, 'rolling ball of 10 pixel radius'
   c) Correct images for varying x-ray contrast levels and different exposure times during scanning
      - Find the mean optical density of the background (MDB); use the average of several measurements over small areas not containing defects
      - Multiply all pixels (background and defects) by 5/MDB
   d) Digitize image: select an area (all or part of the image), and transfer the selection into a matrix containing pixel values
      - Export selection as a data set: *filename.tif*

3) In Transform:
   a) Read *filename.tif*
   b) Define defects: examine each pixel value
      - If value ≤ 17.5, set it to 0 (background)
      - If value > 17.5, set it to 1 (defect)
   c) Remove obvious non-defects (for example, handwriting on right side of x-ray shown in Figure 17)
   d) Save the result as a matrix in text format: *filename.txt*

4) FORTRAN program:
   a) Read *filename.txt* (binary matrix)
   b) Compute the total area and circumference of defects:
      - Defect area $A_{defects}$ = number of defect pixels (pixels with a value of "1")
      - Defect circumference $CF_{defects}$ = number of "1" pixels adjacent (up, down, left, right) to at least one "0" pixel
      - Total area A_{total} = total number of pixels
   c) Defect area percentage = $A_{defects}/A_{total}$
   d) Defect circumference ratio = $CF_{defects}/sqrt(A_{total})$
   e) Equivalent defect radius = $2A_{defects}/CF_{defects}$

The multiplication factor 5/MDB in Scion Image and the cutoff value 17.5 in Transform were selected after considerable trial-and-error. The combination of the constants 5 and 17.5 in these values was found to give the best digital representation of the defects seen in the radiographs. These constants are highly dependent upon each other.

Figures 9 and 10 show the results of steps 1-3 above for reference radiographs CA-1 and CB-5, respectively. These figures demonstrate that the image processing algorithms listed in steps 2 and 3 result in binary digital images that capture the defects on the scanned x-rays quite well. Figure 11 shows a simple example of a binary matrix and the corresponding area and circumference calculations from step 4 above. It should be noted that computing the circumference as shown in the figure underestimates the actual circumference, since this method counts pixels and not pixel edges. During the course of this study, the circumference was also calculated by counting the number of interfaces between adjacent "1" and "0" pixels (i.e. computing the actual circumference of the defect in the binary matrix shown in Figure 11). For the matrix in Figure 11, this method yields a circumference of 24, rather than 13. This method of calculating the circumference overestimates the actual circumference, because while it is calculating the circumference around the "1" pixels, this circumference is longer than the actual value because all edges are horizontal or vertical. The results obtained using these two circumference representations produced the same trends, primarily differing by a scaling factor. To simplify the results, then, only one set of circumference results is given in this paper.

The procedures outlined above were performed on the E186 reference radiographs for shrinkage defect classes CA, CB and CC, Levels 1-5 each. The resulting defect area percentages and defect circumference ratios are plotted against x-ray level in Figures 12 and 13, respectively. The insert plots in these two figures show individual values for the multiple scans of reference radiographs CB1, CB2, CB4 and CB5 mentioned at the beginning of this subsection, along with the line passing through the average values. In both Figures 12 and 13, the individual values for CB1 and CB4 are very close to each other. This indicates that the difference in exposure time (and hence contrast) between the two scans of these radiographs was well accounted for by the Scion Image procedure. However, the two values for CB2 in Figures 12 and 13 exhibit more scatter, which is disappointing. For CB5, one of the three scans has defect area and circumference values that are significantly higher than the other two, indicating that there were more defects in that scan area than in the other two. The average value, in this instance, is probably a good representation of the entire radiograph.

Figure 12 shows that the defect area percentages for classes CA and CB are very similar for all five x-ray levels. This is not surprising, considering the similar veinous nature of both of these defect classes (recall that CA is individual veinous strands of defects, while CB is grouped veinous strands). Class CC (spongy defects) has values close to those of CA and CB for Levels 1-3, but then the CC values become much larger than those of the other classes for Levels 4 and 5. Notice that, for a given x-ray level, the area percentages increase with defect class (CA < CB < CC). When the nature of the three defect classes is taken into account, this seems reasonable. Also, for a given shrinkage class, the defect area percentage increases with x-ray level. This is promising, as it indicates it may be possible to determine x-ray level from defect area percentage for an x-ray with a known shrinkage class. However, Figure 12 points out a rather disappointing aspect of the defect area percentage measurements: to use these curves to determine x-ray level, one must know (i.e. subjectively determine) the shrinkage class. It was hoped that, for example, the area percentages for classes CA, CB and CC for a Level 2 x-ray would all be smaller than all area percentages for Level 3, and larger than all area percentages for Level 1. If this were the case, definition of defect type would be unnecessary, and the level could be determined outright. But the only place this clearly occurs is between Levels 3 and 4. There is a small gap between the values for Levels 4 and 5, but the CC4 area percentage is close to the CA5 value.

Figure 13 shows that the same trends seen for defect area percentage are evident for the defect circumference ratio as well. The defect circumference ratio increases with x-ray level for each shrinkage class, and for a given level the circumference generally increases with shrinkage class. As with defect area percentage, the defect circumference ratio may be useful in determining x-ray level if the shrinkage class is known, but the level cannot be determined without this information. Finally, Figures 12 and 13 both illustrate that shrinkage x-ray levels do not seem to vary in a linear fashion over all five x-ray levels.

9

The trend lines in Figures 12 and 13 for class CA and CB defects show that the quantity of defects increases in a somewhat mild, steady fashion between Levels 1 and 4, and then increases sharply between Levels 4 and 5. For class CC defects, there is a steady increase between Levels 1 and 3, then a sharp increase between Levels 3 and 5.

Figure 14 shows the equivalent defect radius for Levels 1-5 of all three defect classes. If one assumes the defects on the x-rays are all circles of the same size, the equivalent defect radius is simply the radius of these circles. Class CA, CB and CC defects are obviously not circular, but the equivalent defect radius still provides insight into the size of the defects, and the relative changes in size with x-ray level. Figure 14 demonstrates that, while there is some increase in defect size from X-ray Level 1 to 5, it is not a large increase. It is also evident that the size does not increase with every level, since the equivalent defect radius decreases for some levels. Because the average defect size does not significantly increase with x-ray level, it can be concluded that increases in shrinkage severity are mainly due to an increase in the number of defects, rather than the size. The range of the equivalent defect radii shown in Figure 14 is about 0.6-1.1 mm. Multiplying this value by two to compute an equivalent diameter, the size of class CA, CB and CC defects for all x-ray levels is on the order of one to two millimeters.

It is noteworthy that the relatively constant size of shrinkage defects over all x-ray levels indicates that defect optical density measurements would not be useful for determining x-ray level. If it was determined that the size of defects increased with x-ray level, then the optical density of the defects would increase as well. However, since the defect size does not increase appreciably, the optical density of defects would not significantly change either. It is stated in the ASTM x-ray rating standards (ASTM E186, E446 and E280) that radiographic defect density compared to background density should not be used to determine x-ray level, because this variable is highly dependent on "technical factors." From this work, it appears that relative defect density would not be a good measure of x-ray level in any case.

3.2 Quantitative Determination of X-ray Level for Production Radiographs

Once the correlations shown in Figures 12 and 13 were established, the ability to determine ASTM shrinkage x-ray level using these relationships was tested. As indicated in the last subsection, it was determined that these correlations could not be used to determine x-ray level without knowledge of shrinkage type. So three production radiographs with known shrinkage type ratings from the first SFSA / University of Iowa steel plate casting trials [5,6] were selected for this part of the study. The radiographs were scanned at the Iowa State University CNDE, and then processed using the Scion Image and Transform procedures listed earlier. The processed images for these three radiographs are shown in Figure 15 (Trial Case 1), Figure 16 (Trial Case 2) and Figures 17 and 18 (Trial Case 3). The scanned radiograph for Trial Case 3 is included in Figure 17 to illustrate an example of writing on the radiograph (far right side of Figure 17(a)) that was removed during image processing. The light streak in the lower right corner of Figure 17(a) was also removed (see Figure 17(c)). Because these radiographs were used in the reliability and repeatability study detailed in Section 2 of this paper, there were seven different x-ray ratings for each one. The individual ratings, mean x-ray levels and 95% confidence intervals for the three trial cases are listed in Table 2.

Table 2   X-ray ratings for the three trial cases investigated.

Trial Case	Rating 1	Rating 2	Rating 3	Rating 4	Rating 5	Rating 6	Rating 7	Other Defects	Mean X-ray	1-sided 95% CI
1	0	CA1	CB2	CA1	CA2	CA1	CA1	A1, B2	1.14	1.69
2	CB2	CB4	CB5	CC2	CB2	CA3	CB3	B2	3.00	2.83
3	CB5	CB5	CB5	CB5	CB5	CA5	CA5	none	5.00	0.00

In Figures 15, 16 and 18, there are several alternate selection areas shown in addition to the complete processed image of the x-ray shown in Figures 15(b), 16(b) and 17(c). Several selection areas were analyzed to determine the impact of the choice of this area on the resulting x-ray rating. The area A_{total} is included in the defect area and circumference measurement variables to accommodate the prorating requirement in the ASTM x-ray rating standards (ASTM E186, E446 and E280). These standards state

that the "area of interest" on a production radiograph should be prorated to the size of the applicable reference radiograph. By forming the ratios $A_{defects}/A_{total}$ and $CF_{defects}/\sqrt{A_{total}}$, this prorating is done automatically. But what exactly is the "area of interest"? It obviously needs to include the defects of interest, but how much sound area around the defects should be included? This is not a big issue in Trial Case 1 (Figure 15), because the defects in a Level 1 x-ray are few in number and relatively spread out. It becomes more of an issue for Trial Case 2 (Figure 16), and is very important for the centerline shrinkage shown in Trial Case 3 (Figures 17 and 18). By computing the x-ray level for several choices of selection area, it will become evident how important this choice is.

The defect area percentage and defect circumference ratio were calculated for all of the selection areas for Trial Cases 1, 2 and 3. These values are plotted against the selection area in Figure 19 for defect area percentage, and in Figure 20 for defect circumference ratio. In both of these figures, it is seen that there is more scatter in the computed quantities for Trial Case 2 than for Trial Case 1, and significantly more scatter for Trial Case 3 than for Trial Case 2. The rightmost symbol for each trial case in Figures 19 and 20 corresponds to the whole processed image (largest selection area). The pixel labels given for Trial Case 3 in Figures 17 and 18 are included in Figures 19 and 20 to identify the selection areas for this case.

The defect area percentages for all the selection areas shown in Figure 19 were then plotted along the trend line shown in Figure 12 that correlates the defect area percentages of the class CB reference radiographs to their x-ray level. This resulted in the desired relationship between the defect area percentages of the trial case selection areas and ASTM x-ray level. This relationship is shown in Figure 21. An analogous procedure was performed for the defect circumference ratios shown in Figure 20 and the class CB trend line shown in Figure 13, and the resulting plot is given in Figure 22. In Figures 21 and 22, the individual selection areas for each trial case are represented by hollow symbols. The filled symbols indicate the mean value of all the selection areas for each trial case. The error bars for the filled symbols correspond to a 95% confidence interval for x-ray level.

As shown in Table 2, Trial Cases 2 and 3 were predominantly assigned type CB shrinkage (see Table 2), and Trial Case 1 was predominantly CA. For simplicity, all three trial cases were compared to the CB reference radiograph trend line. This was considered acceptable because Trial Case 1 is a nearly sound x-ray, and there is very little difference between CA and CB shrinkage at X-ray Level 1. This is evidenced by the similarity in the CA1 and CB1 values in Figures 12 and 13. Using the CA1 value instead of the CB1 value in Figures 21 and 22 would not change the resulting x-ray level for Trial Case 1. Furthermore, Trial Case 1 was given a CB rating by one radiographer.

Figures 21 and 22 show that all the Trial Case 1 selection areas yield very similar values. All of the selection areas have defect area percentages and defect circumference ratios that are below the corresponding values for X-ray Level 1. The ASTM x-ray rating standards state that one should assign an x-ray level by rounding up to the next shrinkage severity level greater than that of the production x-ray. Therefore, both figures indicate that Trial Case 1 is X-ray Level 1. This agrees with the mean level of 1.1 assigned to this x-ray. For Trial Case 2, Figure 22 shows that all of the selection areas give defect circumference ratios between the Level 1 and Level 2 values, indicating that this x-ray is Level 2. However, the defect area percentages shown in Figure 21 for Trial Case 2 fall both above and below the Level 2 value. The two selection areas below the Level 2 value would indicate that Trial Case 2 is Level 2, and the remaining four areas would indicate Level 3. Viewing the results for Trial Case 2 in both Figures 21 and 22, a case can be made that this x-ray is either Level 2 or Level 3. The conservative approach would be to use Level 3. The mean level for this x-ray is 3.0, but there is considerable variability in the individual ratings (see Table 2). So either a Level 2 or a Level 3 rating would not be unreasonable. For Trial Case 3, some of the defect area percentages in Figure 21 are below the Level 5 reference value, and some are above. By ASTM standards, all of these would indicate that Trial Case 3 is Level 5. The agreement on x-ray level is not quite unanimous in Figure 22. There is one defect circumference ratio that falls below the Level 4 reference value. In fact, it is just below the Level 3 value (see the hollow circle on the left side of the Level 3 reference diamond in Figure 22). This circle corresponds to the (261 x 210) selection area shown in Figure 18. Only part of the centerline shrinkage is contained in this area, and it is mostly one connected defect. Thus, the defect circumference in this

selection area is smaller than in the other areas chosen. But it is clear from the defect circumference ratios of the remaining selection areas, and from the defect area percentage results, that Trial Case 3 is indeed Level 5. This agrees with the mean level of 5.0.

The analysis of the three trial cases presented here demonstrates that a quantitative determination of ASTM x-ray shrinkage level may be possible, provided that the shrinkage type is correctly determined. From the examples shown here, it is evident that the choice of selection area can be very important. The use of multiple selection areas in this study provided additional data points to help determine the level. Based on these results, it seems that both the defect area percentage and defect circumference ratio should be considered when assigning an x-ray level to a radiograph. It must be re-emphasized that the correct selection of shrinkage defect class (CA, CB, CC) is critical for this type of x-ray analysis, since different classes have different reference values used to determine x-ray level. Finally, it seems that the same type of quantitative analysis could also be performed for Category A (gas porosity) and B (inclusion) defects, due to the similar nature of their current rating structures.

4. Summary and Conclusions

Two different studies were presented in Sections 2 and 3 that examined the ASTM x-ray shrinkage rating system for radiographs of steel castings. The first study evaluated the reliability and repeatability of x-ray shrinkage ratings through a statistical study. The second study involved an effort to determine the shrinkage severity level of x-rays through digital analysis of scanned radiographs.

The reliability and repeatability of ASTM shrinkage x-ray ratings was investigated in a statistical study performed on 128 x-rays, each of which were rated seven different times. It was found that the seven ratings were in unanimous agreement on shrinkage type for 37% of the x-rays, on shrinkage level for 17% of the x-rays, and on both type and level for 12.5% of the x-rays. The x-rays that had unanimous agreement on both type and level were all either completely sound, or very unsound (Level 5). By computing a 95% confidence interval for the seven shrinkage level ratings for each x-ray, it was found that x-rays with an average level around 2 or 3 had the largest variability in level ratings, with an average variability of about ± 2. The different level ratings for each x-ray were then examined to determine how many x-rays had level ratings both above and below common accept/reject thresholds. 21-35% of the x-rays had ratings that crossed different thresholds, indicating that the castings associated with these x-rays would have been accepted by some radiographers and rejected by others. Two of the radiographers involved in this study rated all the x-rays twice, and analysis was performed to determine the consistency of their ratings. Comparing each radiographer's second rating to their first, both radiographers gave different type ratings for at least 19% of the x-rays and different level ratings for at least 34% of the x-rays. Both radiographers also reversed accept/reject decisions for 10-15% of the x-rays. This study indicated the subjective nature of the x-ray rating procedure.

Next, a study was performed to attempt to quantify ASTM shrinkage x-ray rating. ASTM E186 reference radiographs for shrinkage classes CA, CB and CC were scanned and digitized, and measurements of the defects were performed. It was found that the defect area percentage and defect circumference ratio both increased with x-ray level for a given defect class. Also, for a given x-ray level, these measures generally increased with defect class (CA < CB < CC). The increase of defect area and circumference with x-ray level for a given defect class indicates the possibility of determining x-ray level with these measures. However, it was determined that these defect measurements could not be used to determine x-ray level unless the shrinkage type was known. This study also indicated that defect size does not increase considerably with x-ray level. Rather, severity increases due to an increase in the number of defects. The defect size for all three shrinkage classes studied is on the order of one to two millimeters. Once the reference radiographs had been thoroughly analyzed, three production x-rays were digitized. Defect area percentages and defect circumference ratios were computed for various selection areas of each x-ray. By comparing these values to the corresponding class CB shrinkage reference radiograph values, it was possible to correctly determine the x-ray level. It was found, however, that different selection areas could change the resulting x-ray level. Use of multiple selection areas and both the defect area and circumference measures resulted in reasonable x-ray levels. The results of this study indicate that, through digital x-ray analysis, it is possible to reduce the uncertainty in ASTM shrinkage x-

ray level rating. However, this digital analysis cannot be done until the correct shrinkage type rating is assigned to an x-ray.

ACKNOWLEDGMENTS

REFERENCES

[1]   American Society for Testing of Materials, ASTM E 446, "Standard Reference Radiographs for Steel Castings Up to 2 in. (51 mm) in Thickness," *1998 Annual Book of ASTM Standards*, Volume 03.03: Nondestructive Testing, 1998.

[2]   American Society for Testing of Materials, ASTM E 186, "Standard Reference Radiographs for Heavy-Walled (4½ to 12-in. (114 to 305-mm)) Steel Castings," *1998 Annual Book of ASTM Standards*, Volume 03.03: Nondestructive Testing, 1998.

[3]   American Society for Testing of Materials, ASTM E 280, "Standard Reference Radiographs for Heavy-Walled (2 to 4½-in. (51 to 114-mm)) Steel Castings," *1998 Annual Book of ASTM Standards*, Volume 03.03: Nondestructive Testing, 1998.

[4]   Quantrell, R. J., "Evaluation of the Consistency of Radiographic Grading for Shrinkage Defects", *SCRATA Committee Paper*, 1981.

[5]   Hardin, R. A., Shen, X., Gu, J., and Beckermann, C., "Progress in the Development of Improved Feeding Rules for the Risering of Steel Castings," *1998 SFSA Technical and Operating Conference*, 1998.

[6]   Hardin, R. A., Shen, X., Gu, J., and Beckermann, C., "Use of Niyama Criterion to Predict ASTM X-Ray Levels and to Develop Improved Feeding Rules for Steelcasting", *AFS Paper No. 99-062*, presented at the 103rd AFS Casting Congress, 1999.
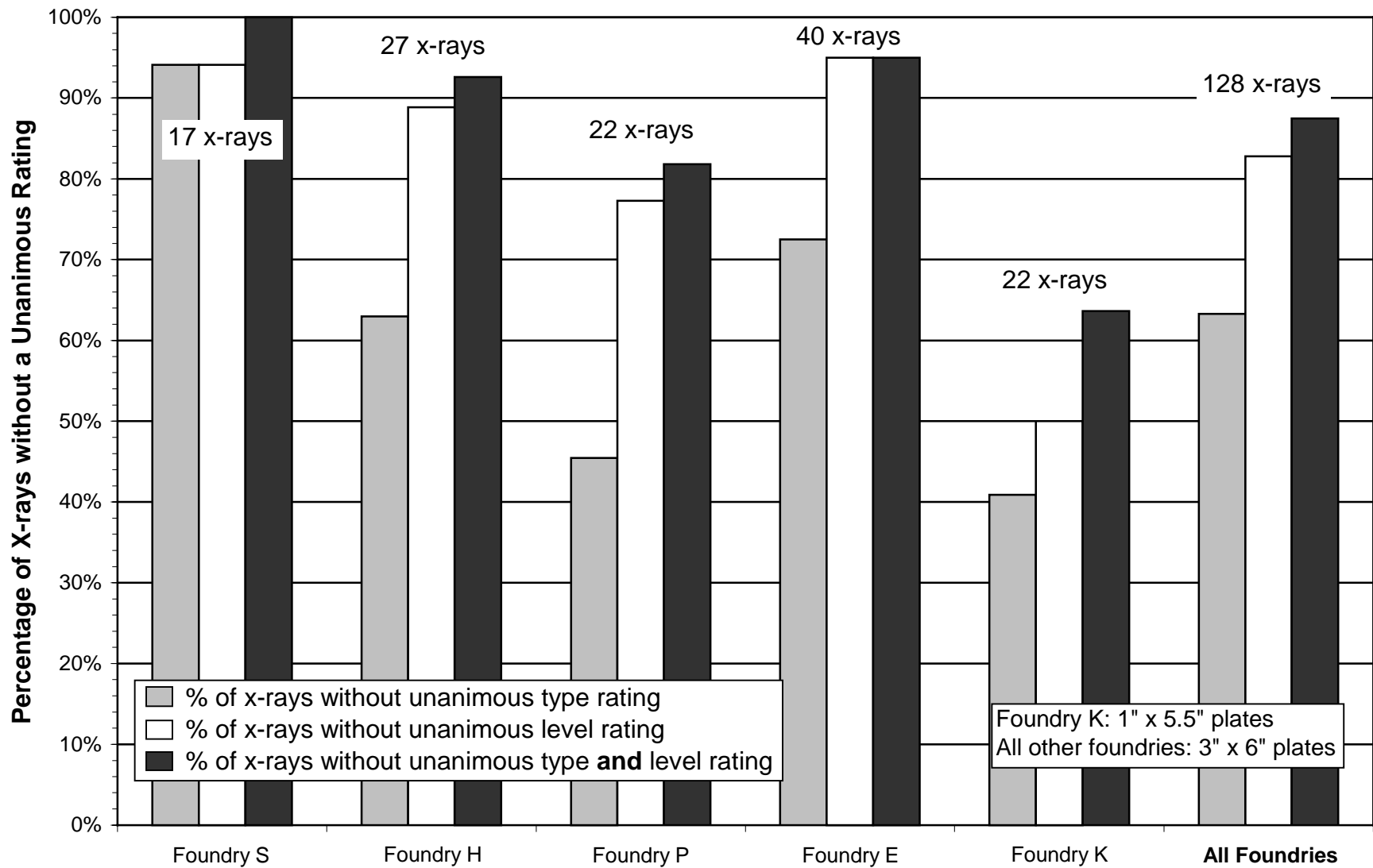
Figure 1   Percentage of x-rays without a unanimous shrinkage type, level, and total rating, grouped by foundry.
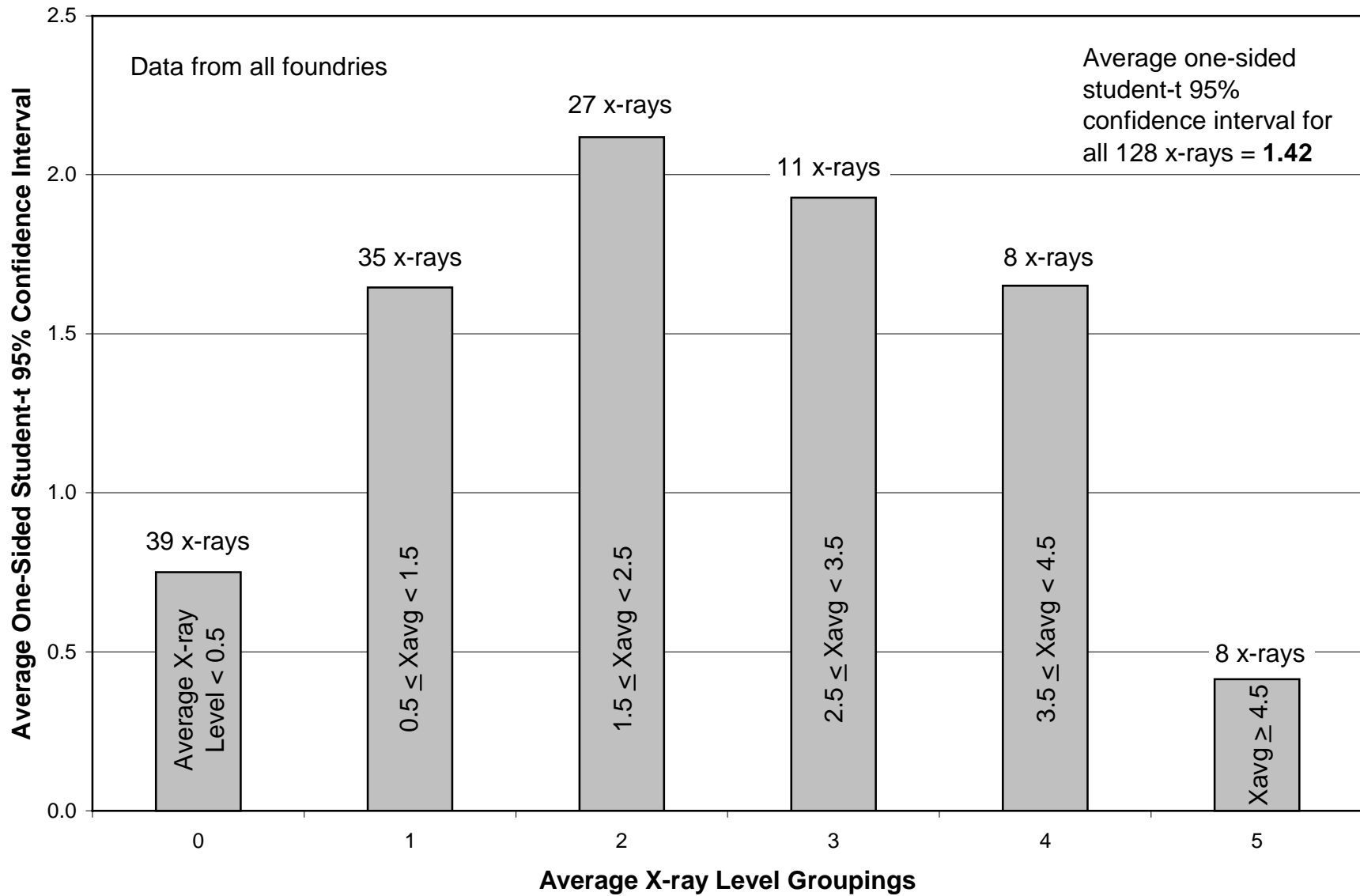
Figure 2   Average one-sided confidence intervals of x-ray level ratings, grouped by average x-ray level.
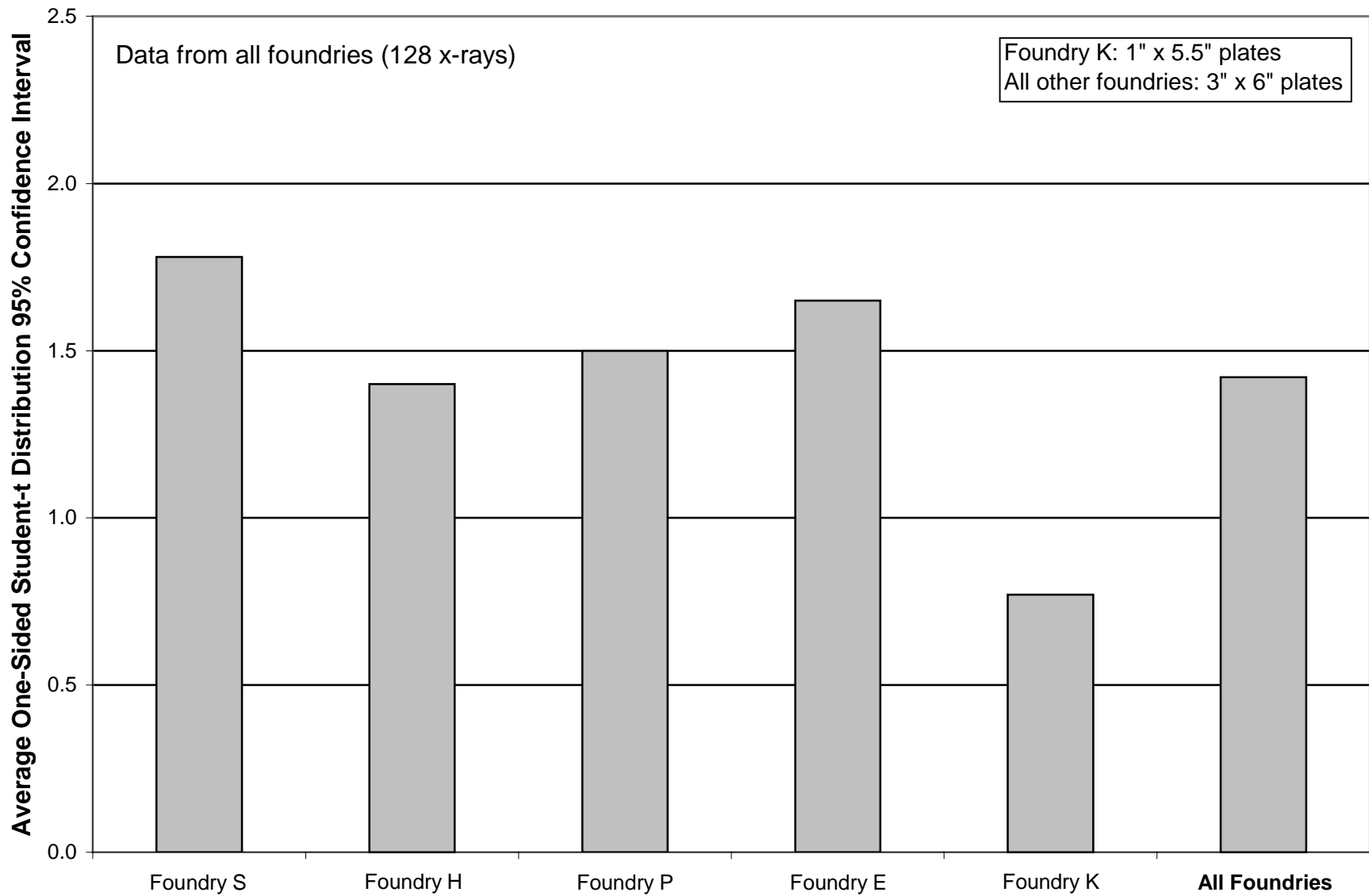
Figure 3   Average one-sided confidence intervals of x-ray level ratings, grouped by foundry.

Number of Shrinkage Level Ratings for an X-ray that Disagree with Consensus
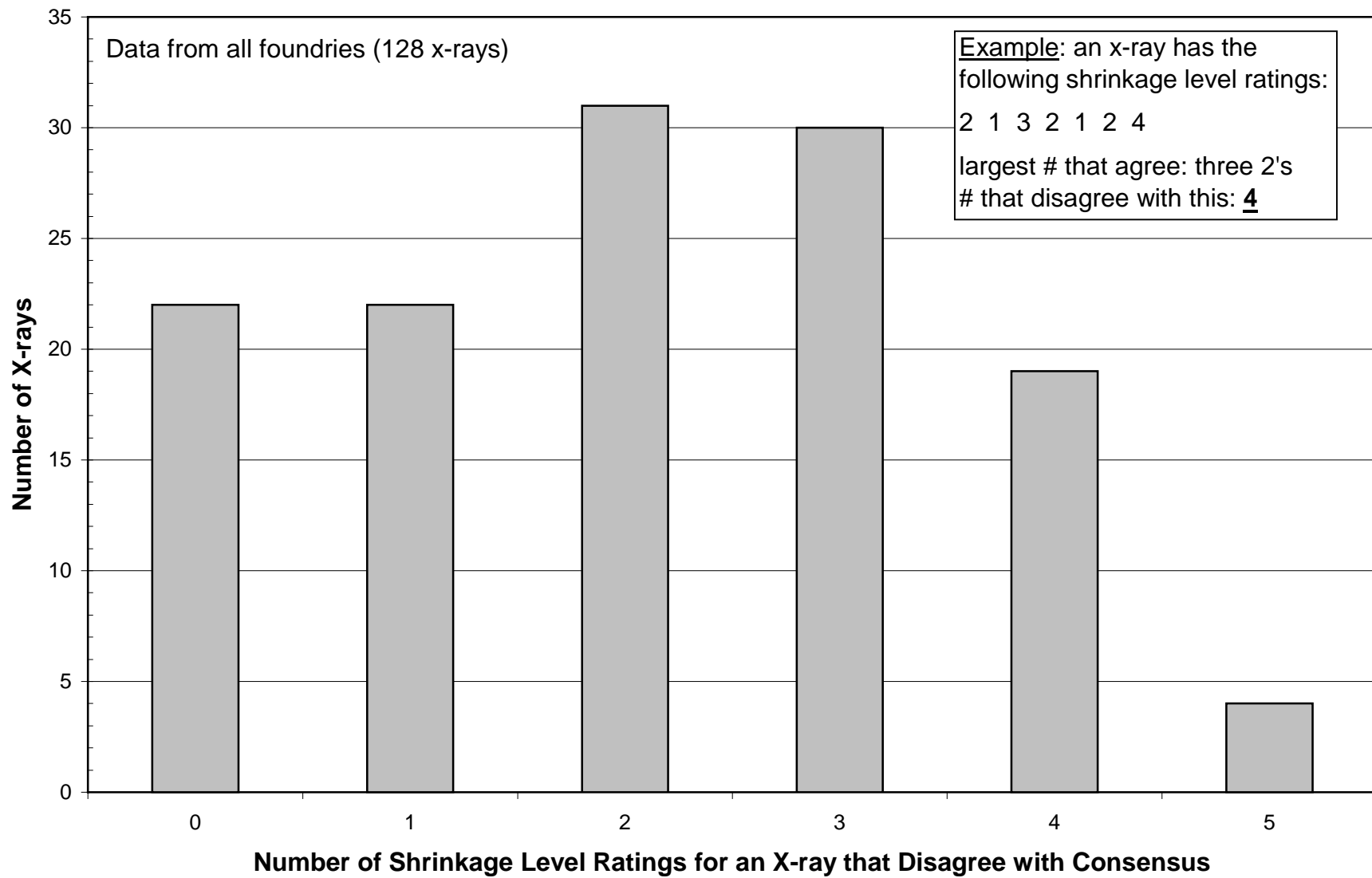
Figure 4   Number of shrinkage level ratings (0, 1, 2, 3, 4, 5) that disagree with the consensus level rating for each x-ray.

Data from all foundries (128 x-rays)

Example: an x-ray has the following shrinkage level ratings:

2 1 3 2 1 2 4

largest # that agree: three 2's
that disagree with this: **4**
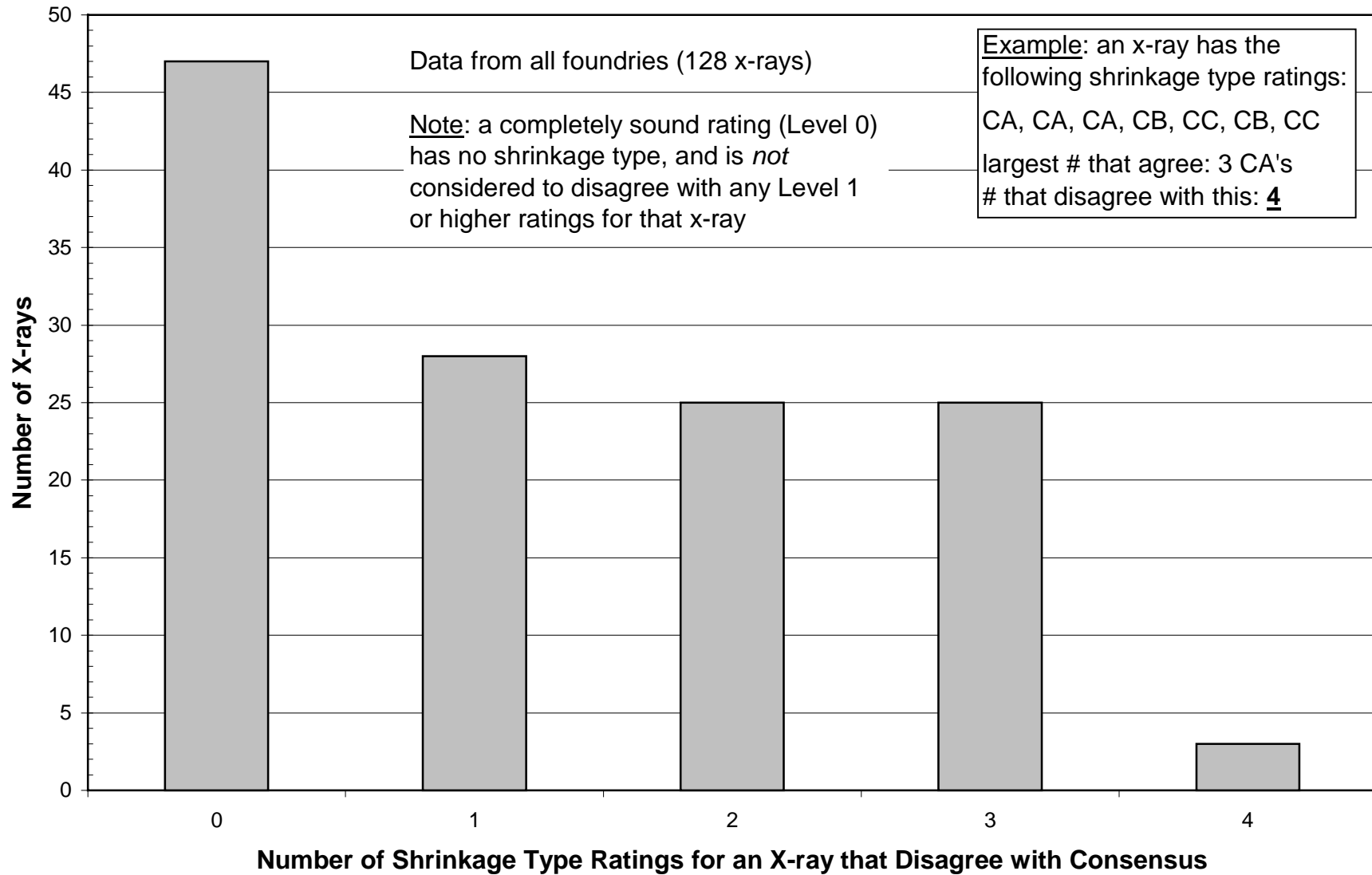
Figure 5   Number of shrinkage type ratings (CA, CB, CC, CD) that disagree with the consensus type rating for each x-ray.
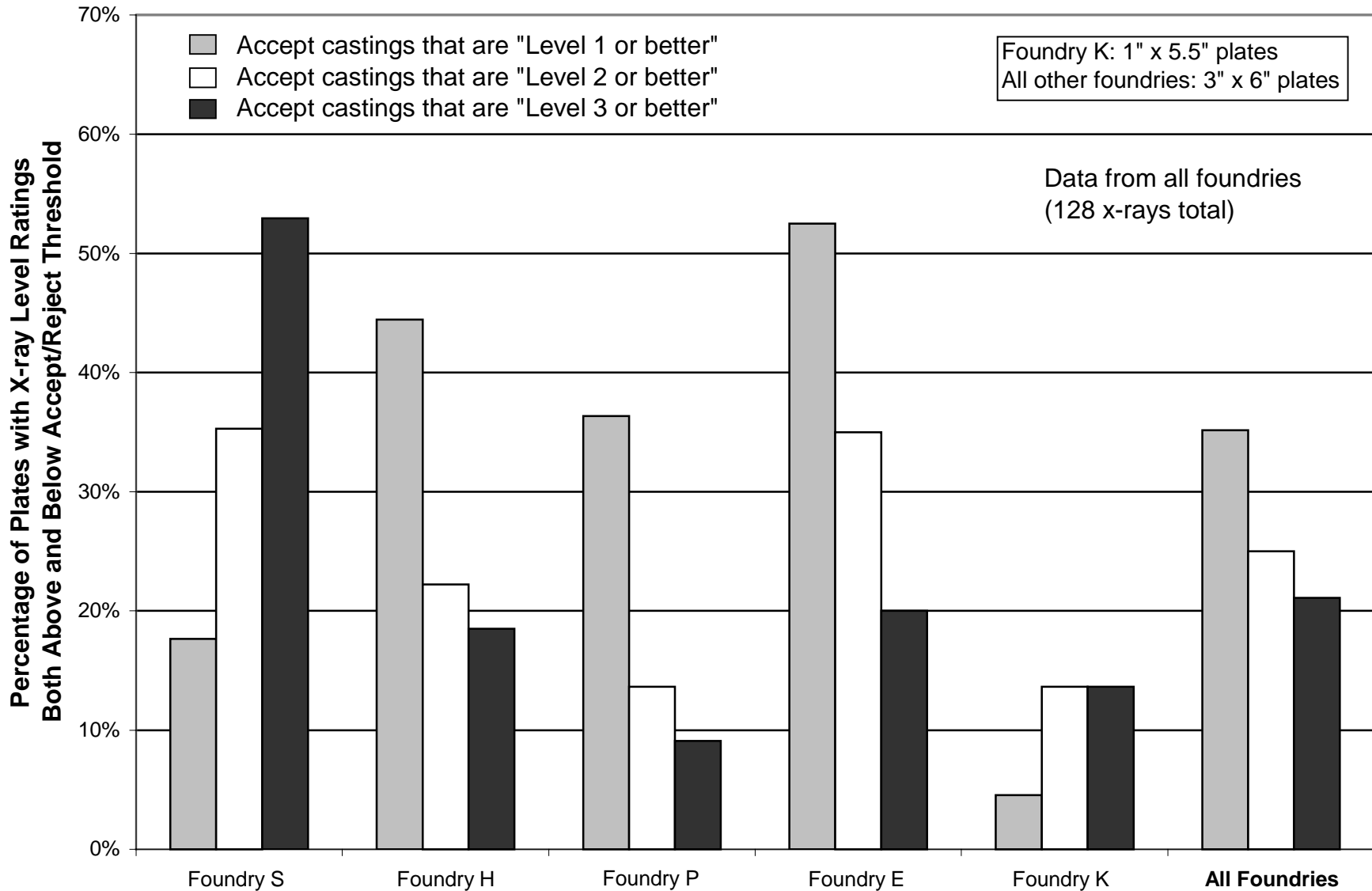
Figure 6   Percentage of x-rays accepted by some radiographers and rejected by others, grouped by foundry.
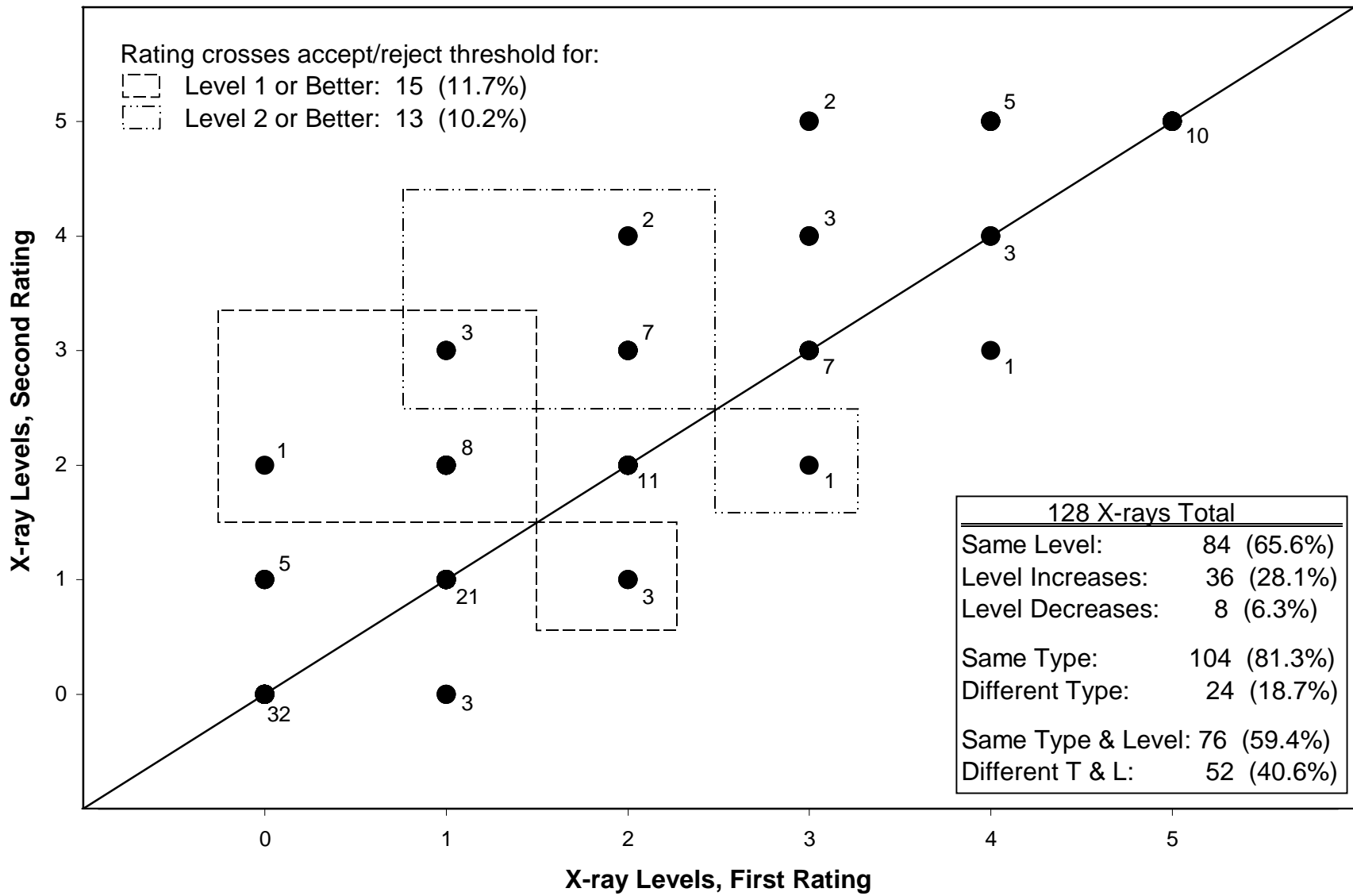
Figure 7   Comparison between first and second ratings of Radiographer 2.
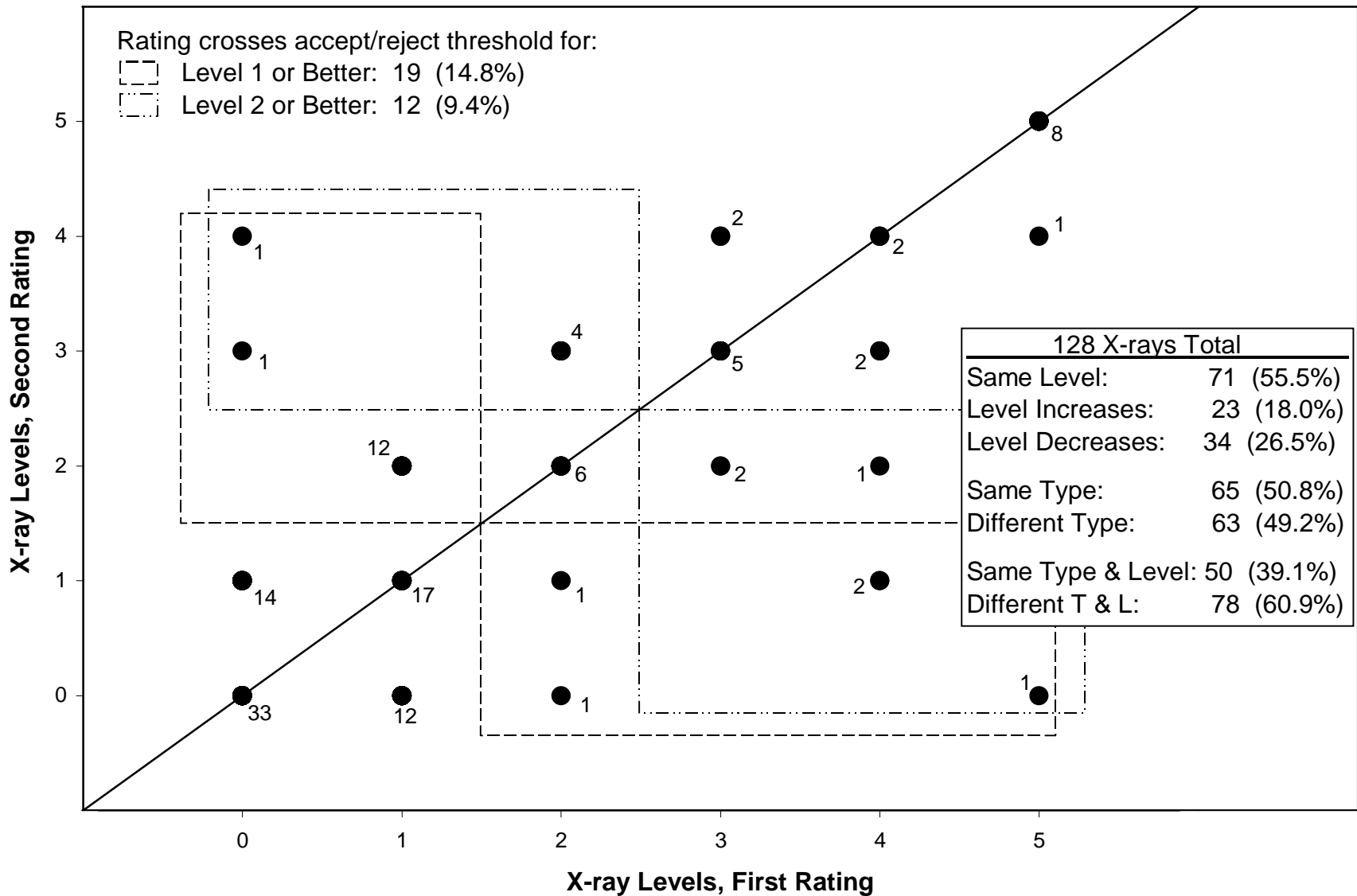
Figure 8   Comparison between first and second ratings of Radiographer 3.
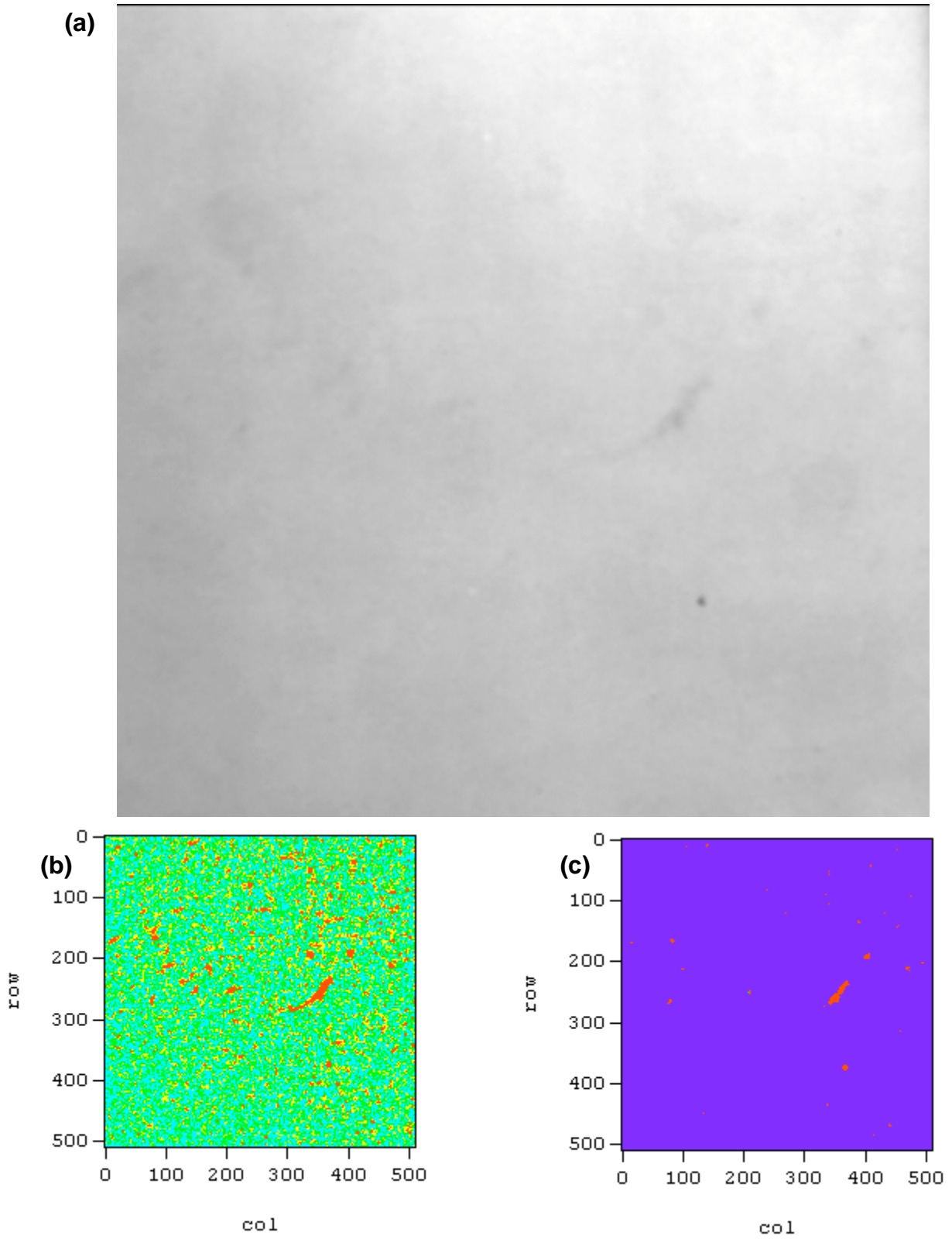
Figure 9   Reference radiograph CA-1 (a) after scanning, (b) after Scion Image processing, and (c) after Transform processing.
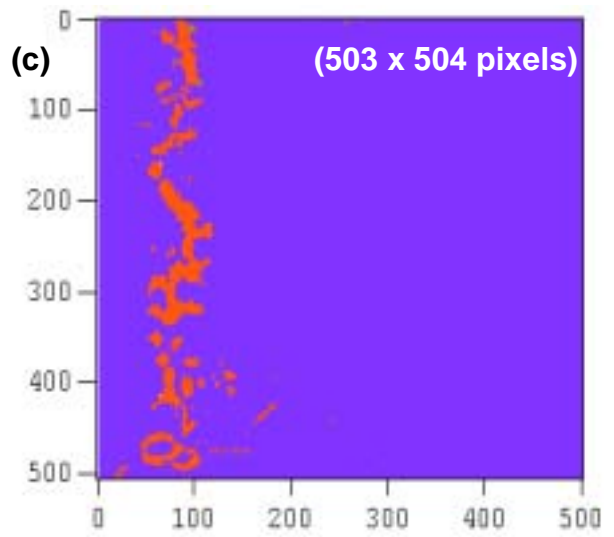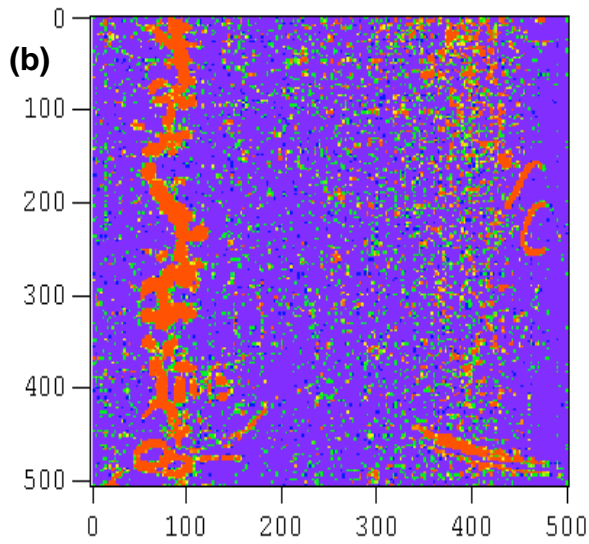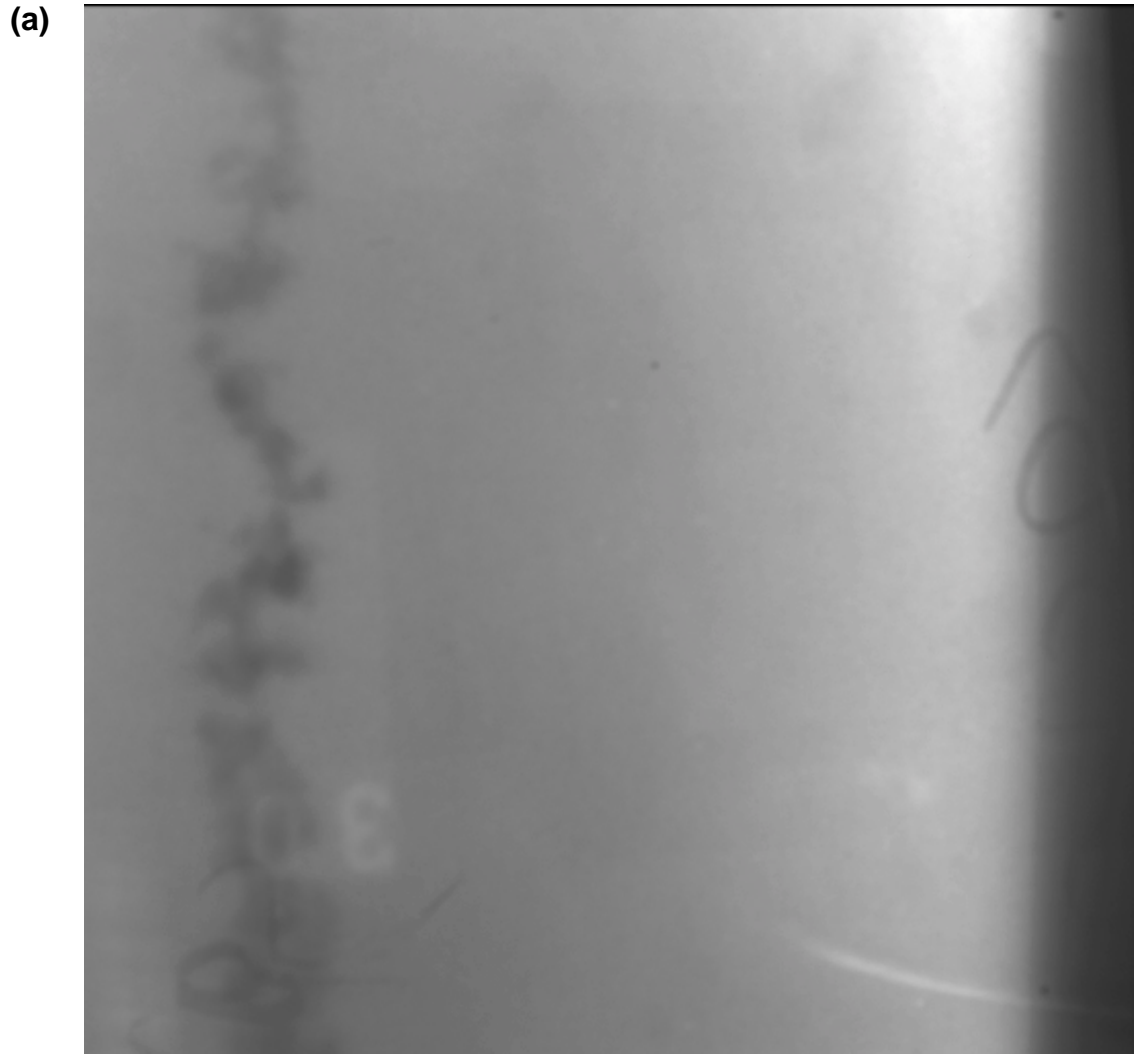
(a)

(b)

(c)
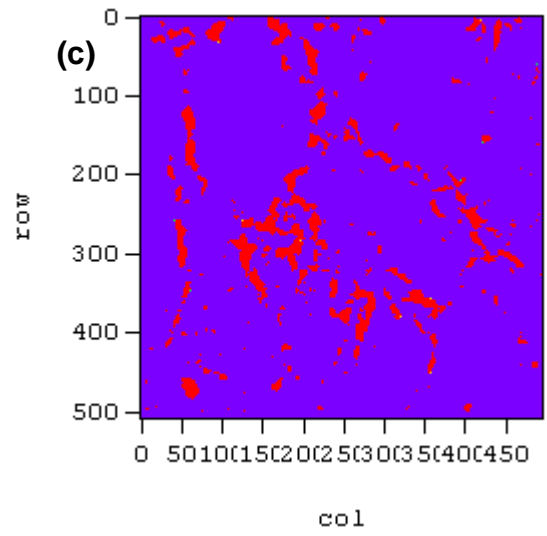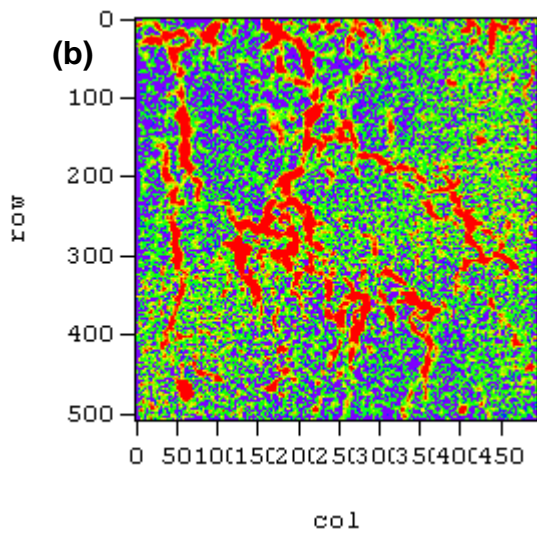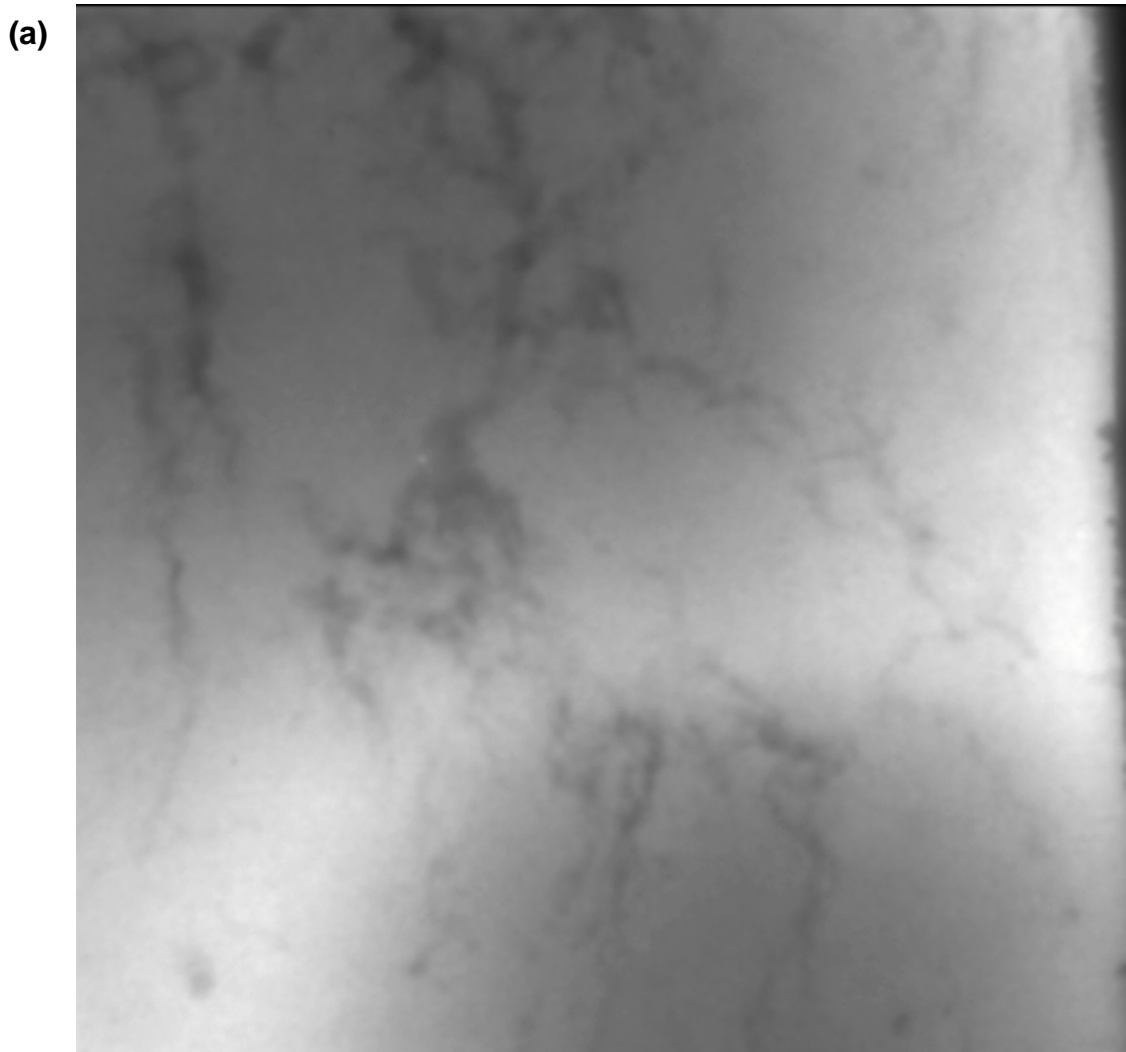
Figure 10   Reference radiograph CB-5 (a) after scanning, (b) after Scion Image processing, and (c) after Transform processing.

0	0	0	0	0	0	0
0	0	0	1	1	0	0
0	0	0	1	1	1	0
0	1	1	1	1	1	0
0	0	1	1	1	0	0
0	0	0	1	1	1	0
0	0	0	0	0	1	0
0	0	0	0	0	0	0

$A_{defects}$ = number of "1" pixels = 17

$CF_{defects}$ = number of "1" pixels adjacent (up, down, left, right) to at least one "0" pixel = 13

A_{total} = total number of pixels = 56

defect area percentage = $A_{defects}/A_{total}$ = 30.4%

defect circumference ratio = $CF_{defects}/sqrt(A_{total})$ = 1.74

Figure 11   Sample binary matrix, with area and circumference calculations.
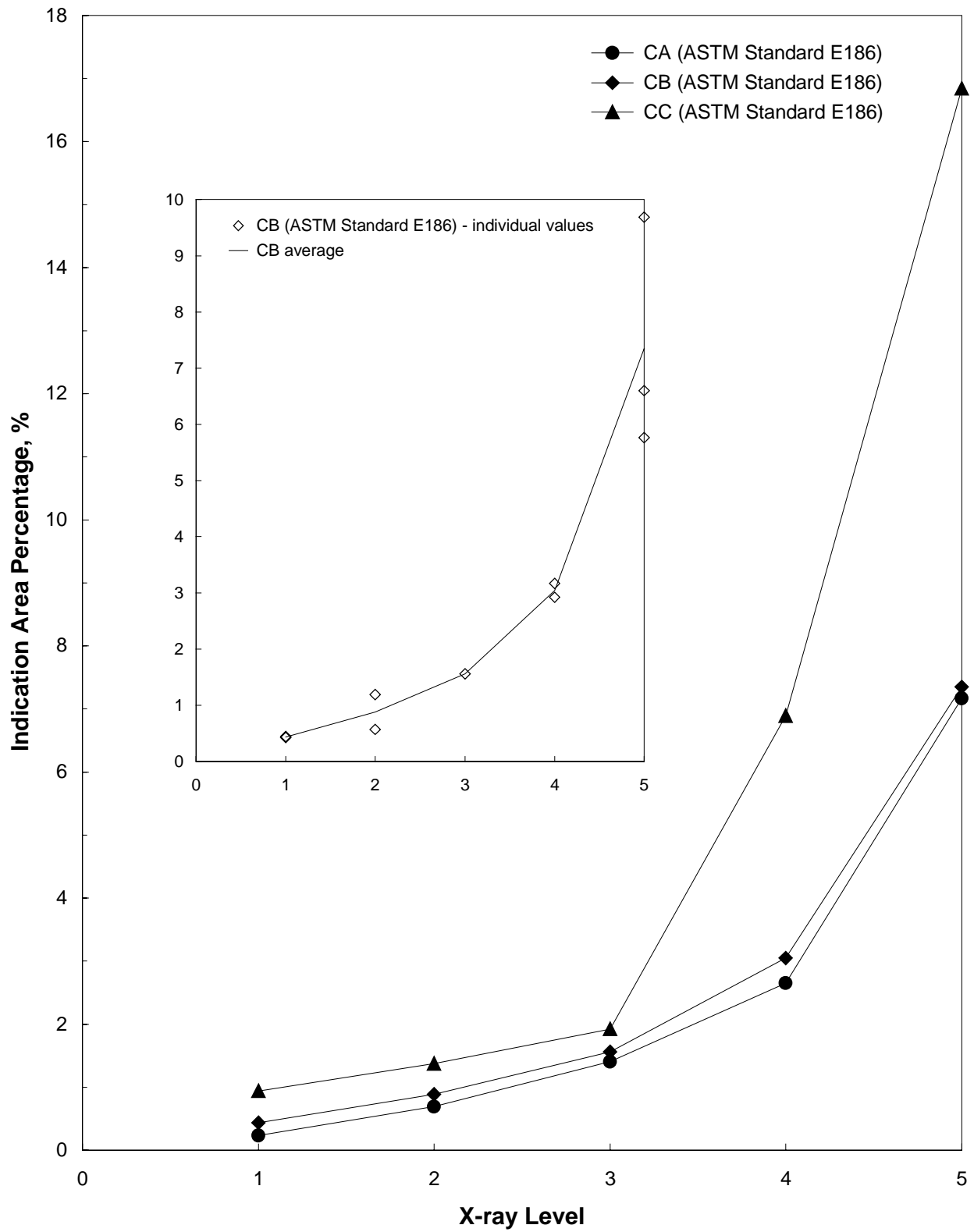
Figure 12   Defect area percentage values for E186 reference radiographs for shrinkage defects (classes CA, CB and CC).
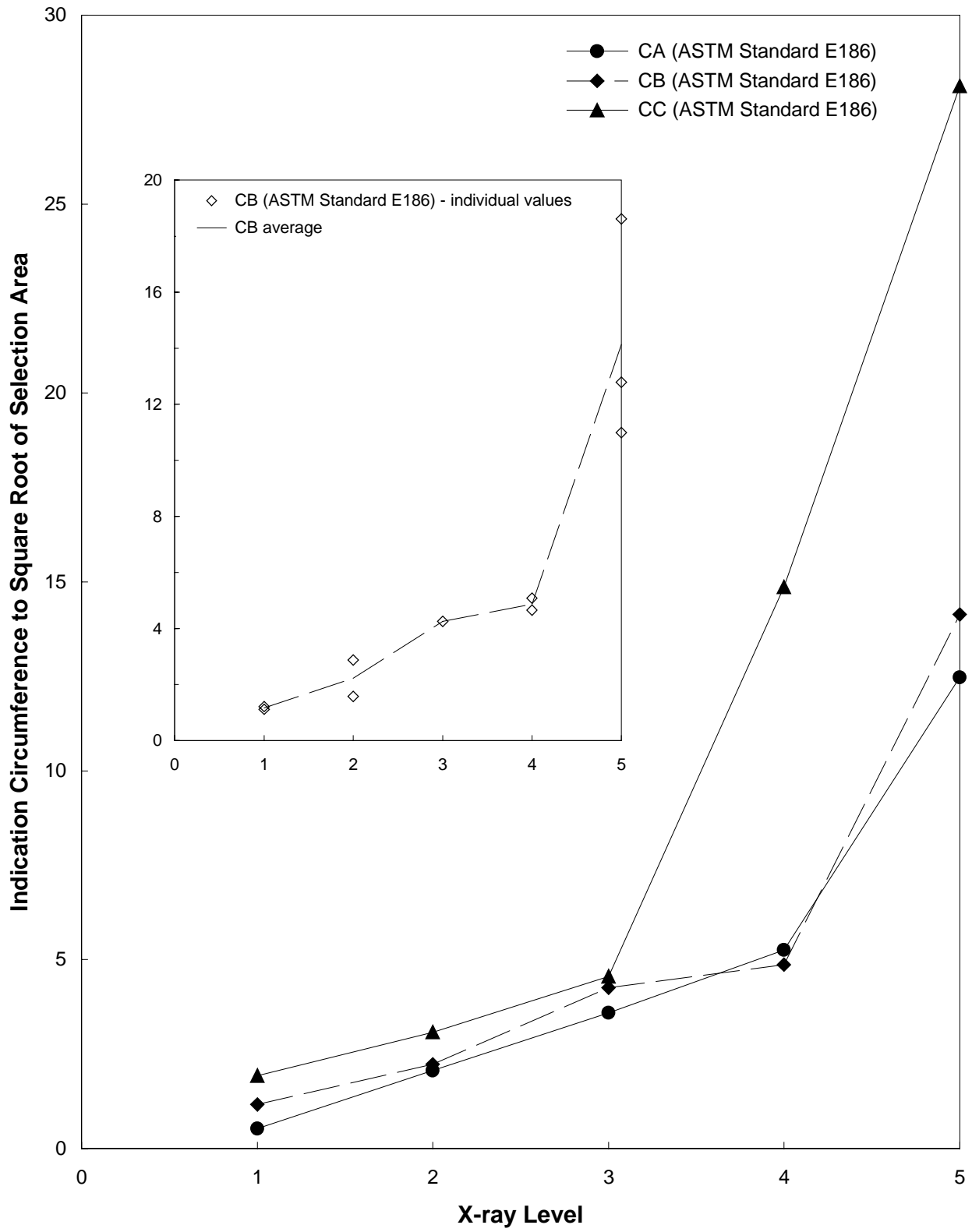
Figure 13 Defect circumference ratio values for E186 reference radiographs for shrinkage defects (classes CA, CB and CC).
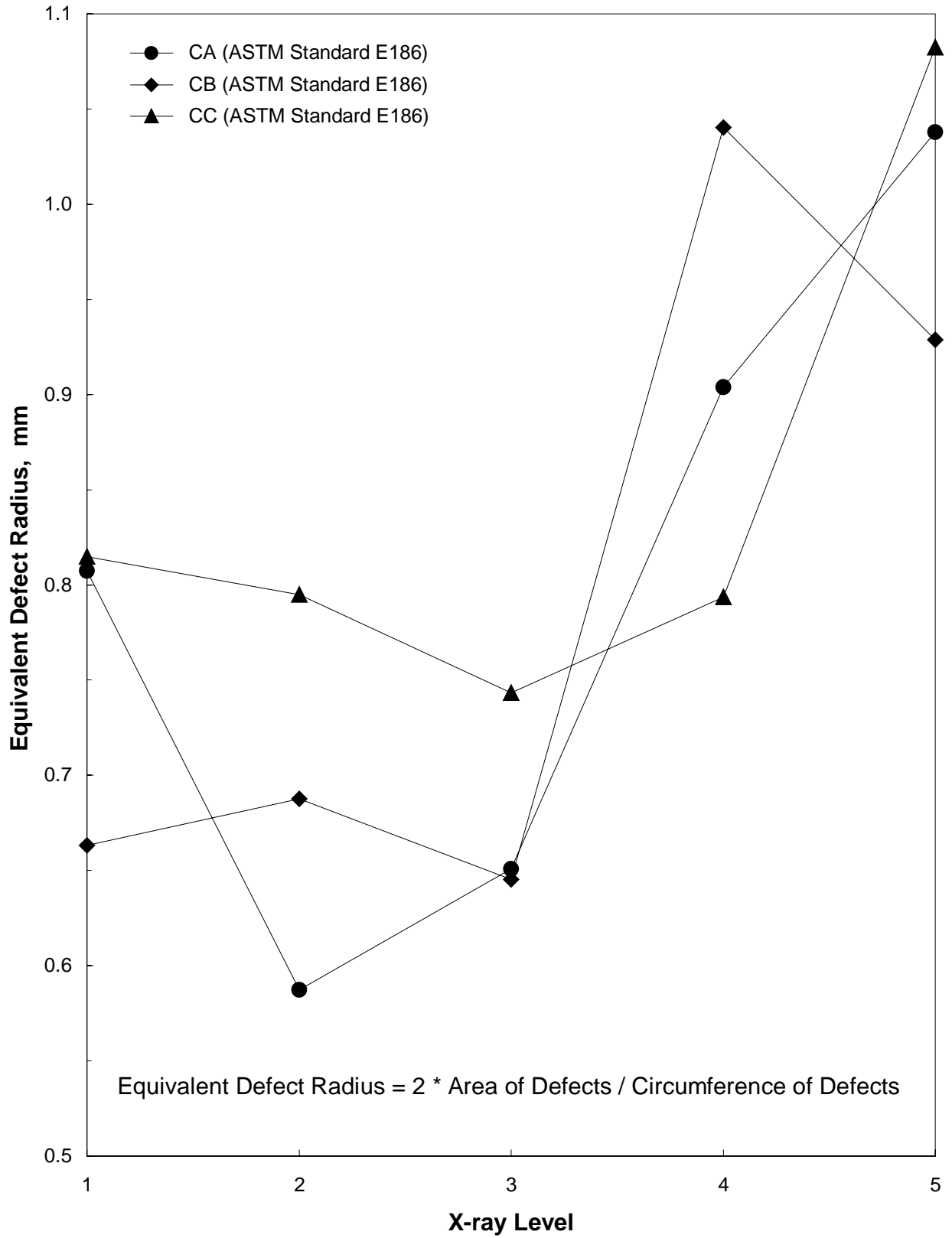
Figure 14 Variation of equivalent defect radius with x-ray level for class CA, CB and CC shrinkage defects.

(a) after Scion Image

(b) after Transform

(494 x 500 pixels)

(c) additional selection areas:

(280 x 176)

(278 x 281)

(272 x 114)

(242 x 492)
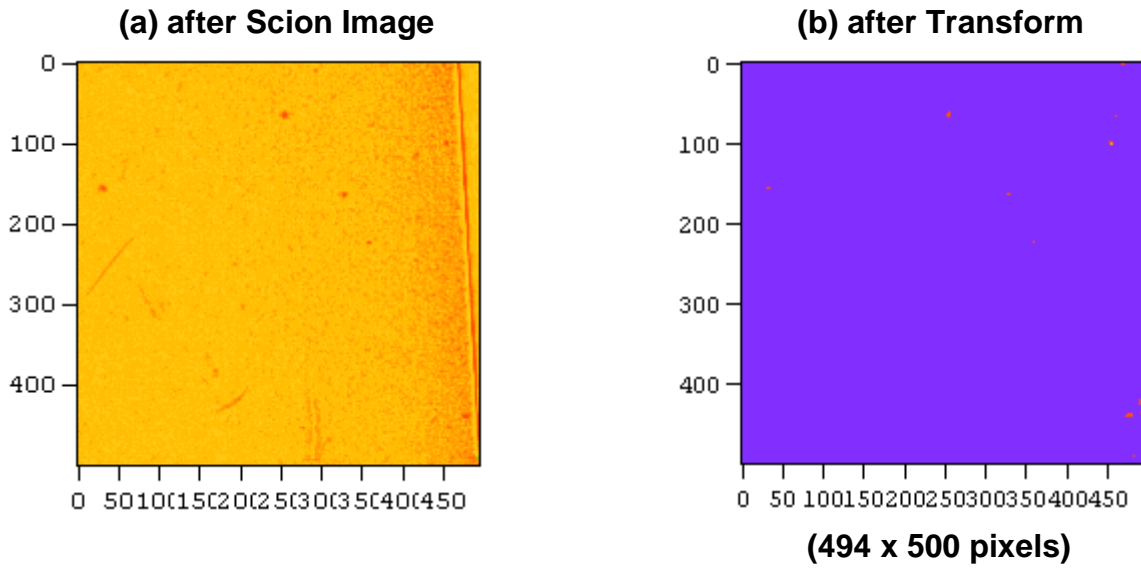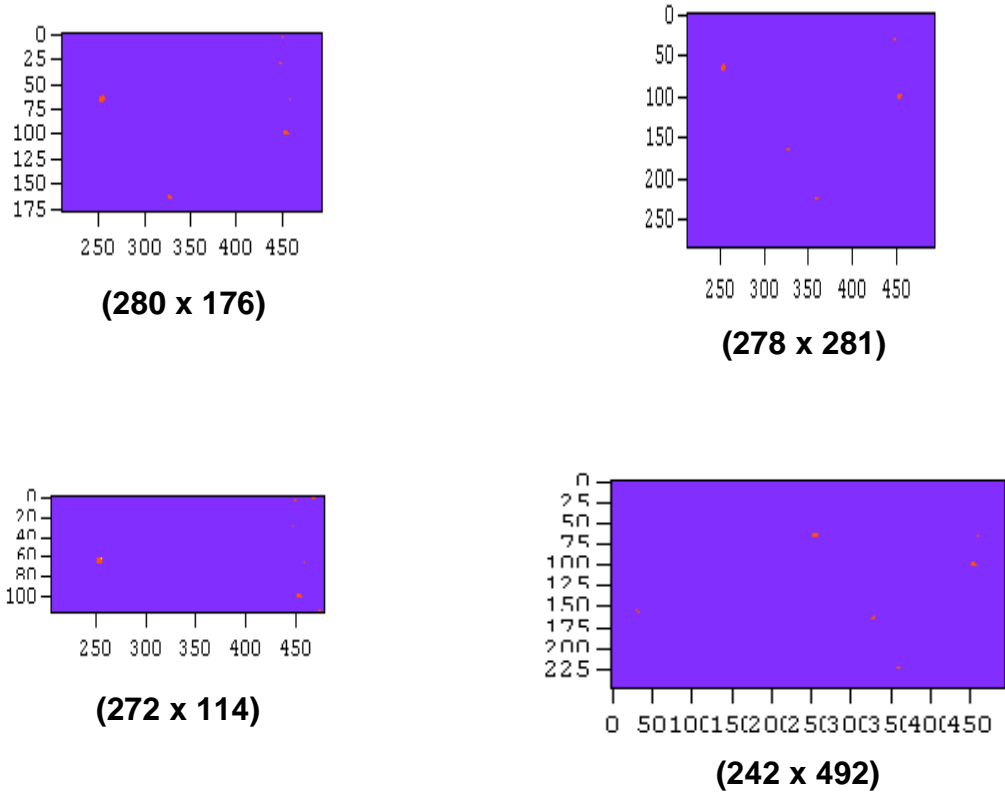
Figure 15   Trial Case 1 (X-ray Level 1) (a) after Scion Image processing, (b) after Transform processing, and (c) several alternate selection areas.

(a) after Scion Image

(b) after Transform

(479 x 503 pixels)

(c) additional selection areas:

(269 x 473)

(449 x 440)

(212 x 373)

(299 x 442)

(195 x 477)
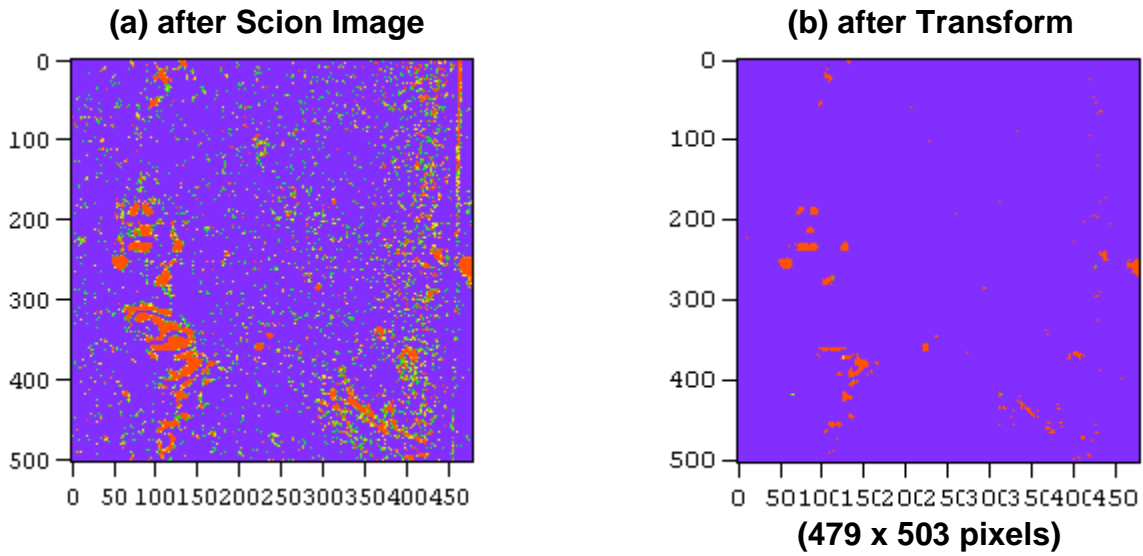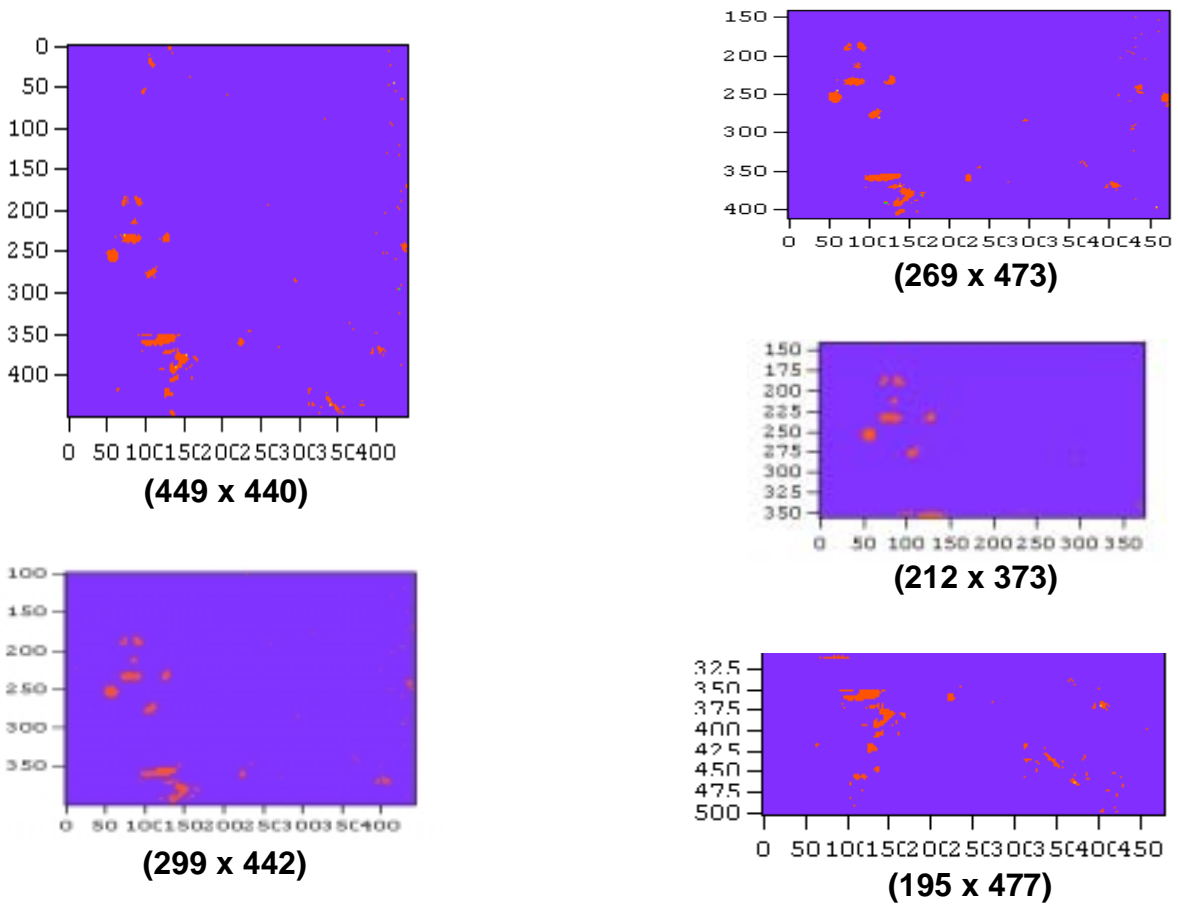
Figure 16   Trial Case 2 (X-ray Level 2) (a) after Scion Image processing, (b) after Transform processing, and (c) several alternate selection areas.

(a)

(b)

(c) (503 x 504 pixels)

Figure 17   Trial Case 3 (X-ray Level 5) (a) after scanning, (b) after Scion Image processing, and (c) after Transform processing.

(170 x 501 pixels)

(134 x 207)

(504 x 235)

(261 x 210)

(496 x 164)
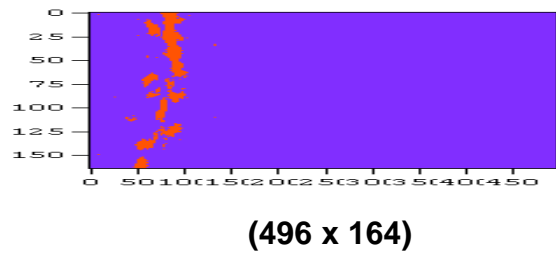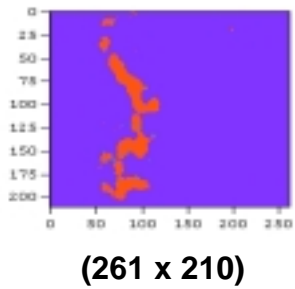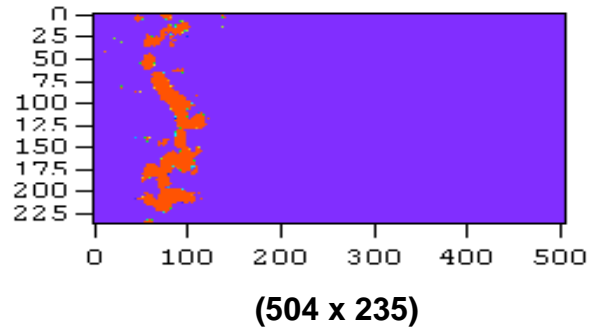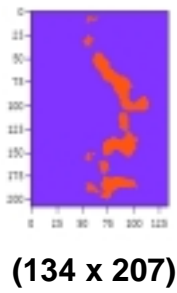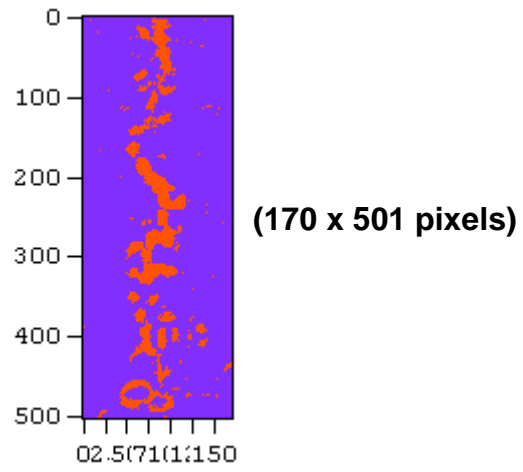
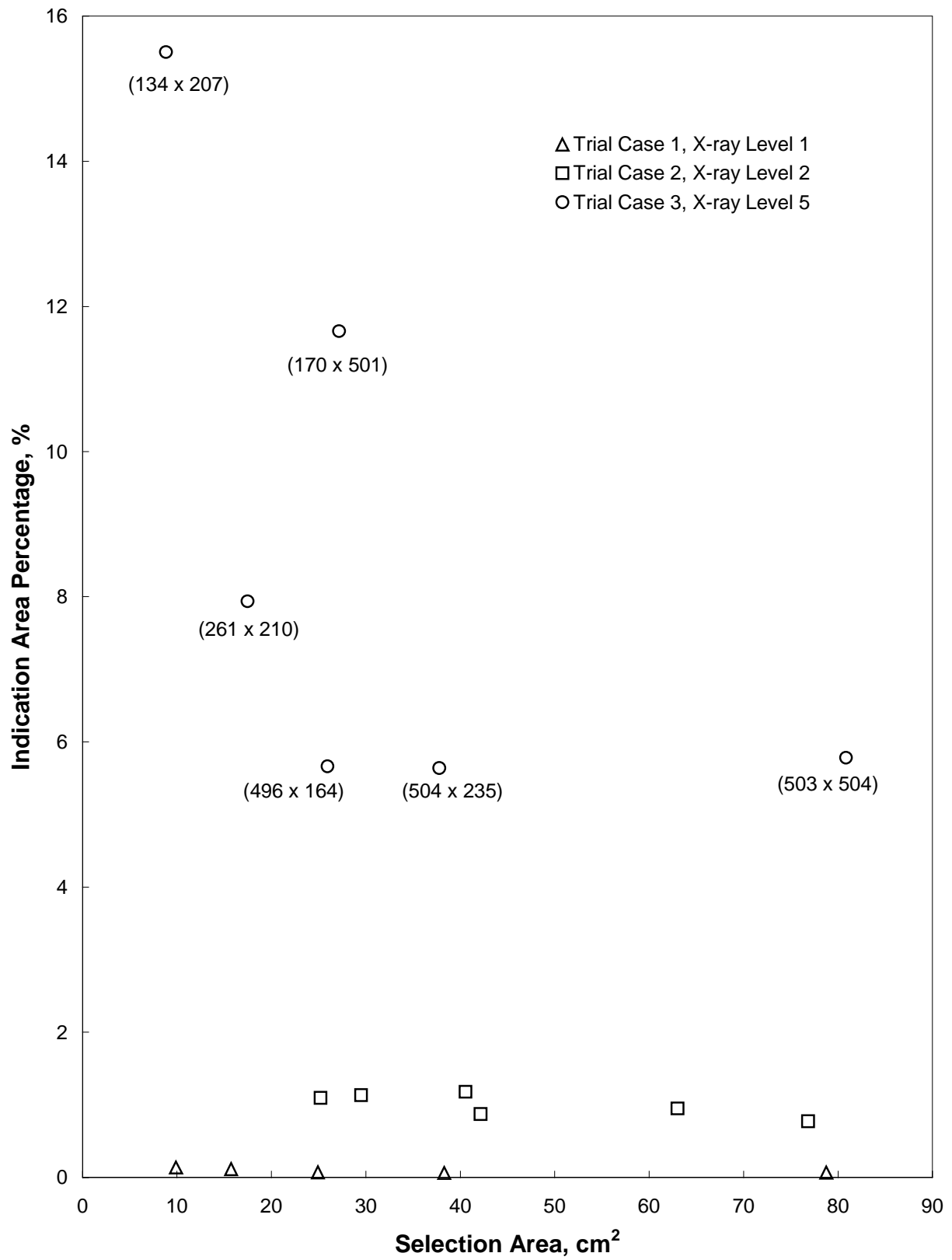Figure 18   Alternate selection areas for Trial Case 3 (X-ray Level 5).

Figure 19 Variation of defect area percentage with choice of selection area for all three trial cases.
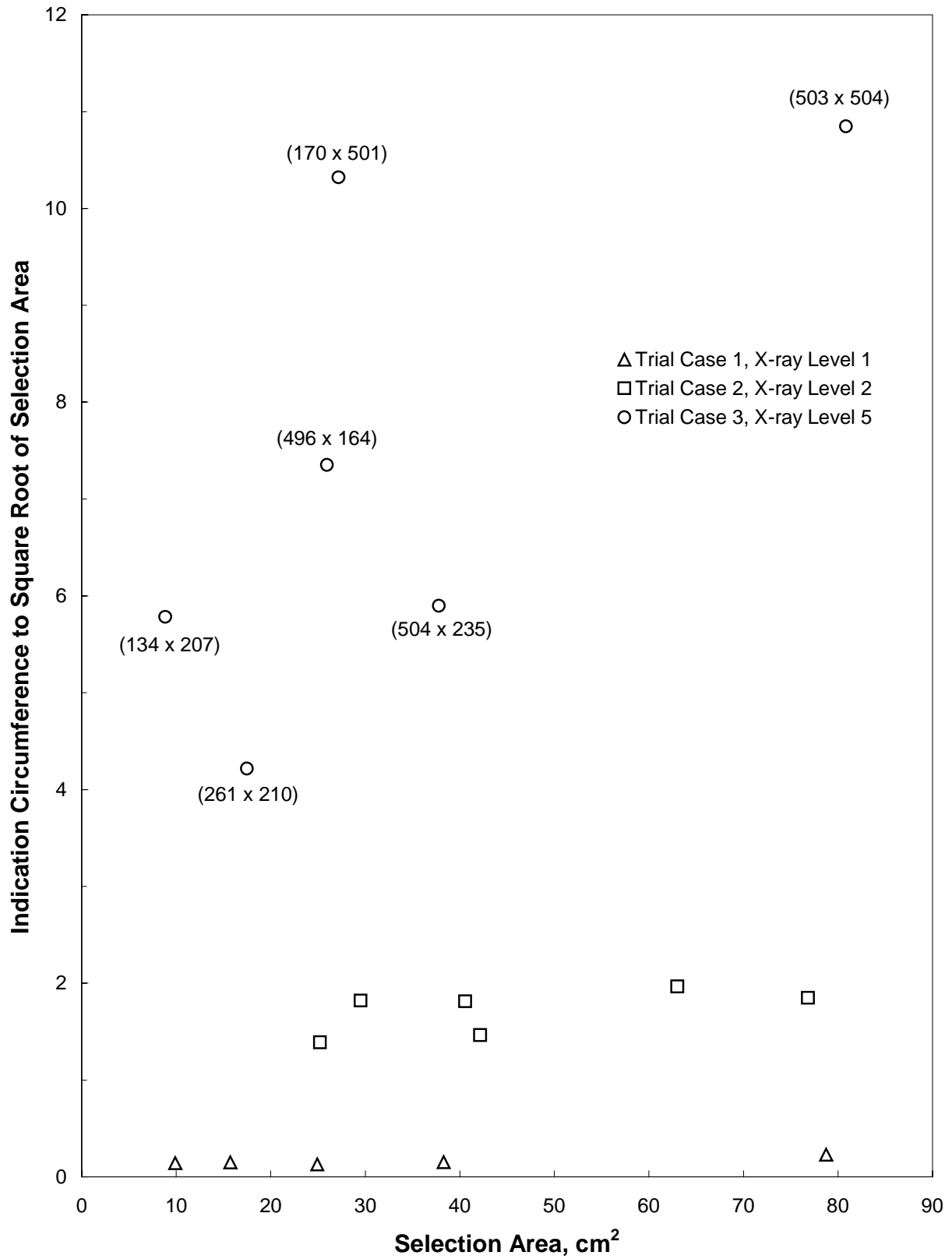
Figure 20 Variation of defect circumference ratio with choice of selection area for all three trial cases.
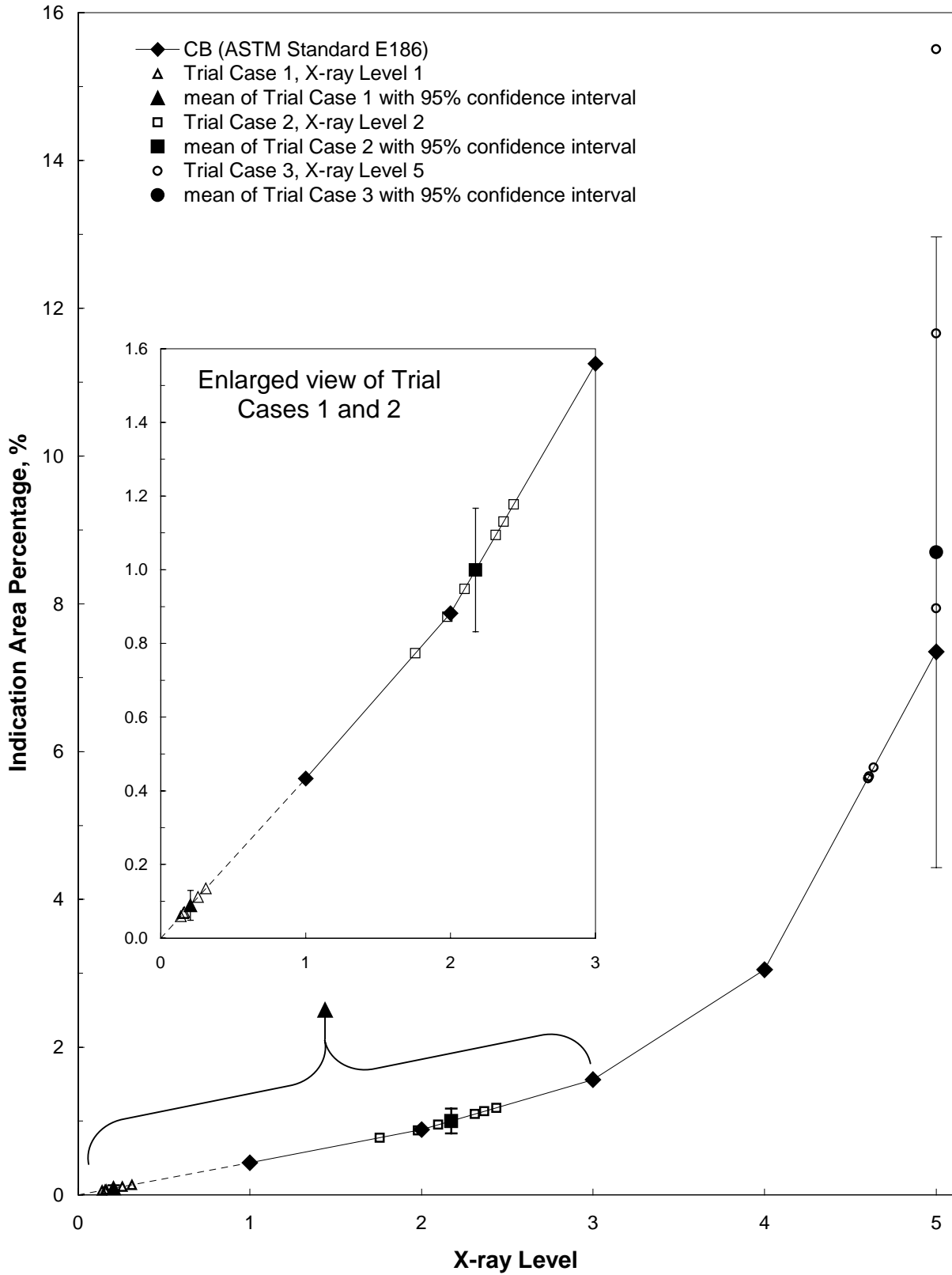
Figure 21   Defect area percentage versus x-ray level for all trial cases, all selection areas.
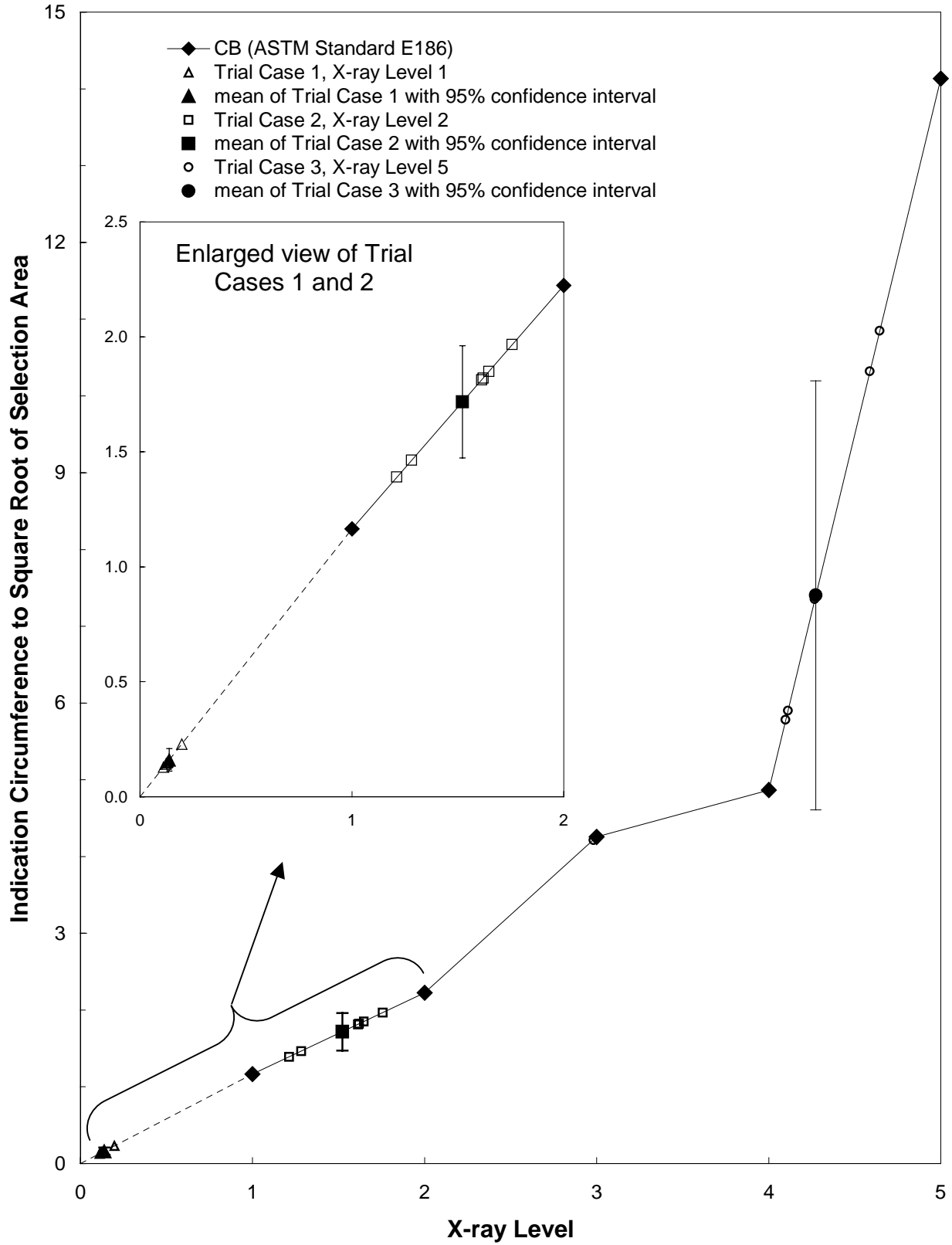
Figure 22 Defect circumference ratio versus x-ray level for all trial cases, all selection areas.