



Standard Setting Report: Breast Sonography - Effective September 2024

Background

The mission of the American Registry of Radiologic Technologists (ARRT) is to “promote high standards of patient care by recognizing qualified individuals in medical imaging, interventional procedures, and radiation therapy.” The ARRT’s equation for excellence states that excellence equals education plus ethics plus examination; standard setting is one of many processes within the examination component that ensure it is an accurate reflection of the knowledge, skills, and abilities required of entry level technologists.

This report details a pair of standard setting meetings conducted in November 2021 and January 2023 for breast sonography, including committee composition, methods, results, recommendations, and any changes to the exam cut score. It is ARRT’s primary goal for the exam to reflect the current state of practice and expectations for entry-level breast sonographers. Therefore, this meeting served to update those expectations from the previous standard setting in 2004.

The ARRT utilizes experts in standard setting, called psychometricians, to train and facilitate a committee of subject matter experts from the field to define expectations, collect data, and make recommendations before presenting the results to the ARRT Board of Trustees. The facilitator provided training throughout the meeting to ensure that the committee was prepared to hold productive discussions, make well-reasoned judgments, and provide suitable recommendations at the meeting’s conclusion.

First Meeting – November 2021

Facilitators:

- Jessica Anderson, ICE-CCP, Director, Test Development, Licensure and Certification at DRC
- Ross Brown, PhD, Senior Psychometrician, Licensure and Certification at DRC

Committee Composition

ARRT staff selected individuals from the volunteer database with the goal of maximizing diversity in role, geography, and experience in breast sonography. When possible, ARRT will bias the volunteer pool towards individuals early in their career as the exam is designed to assess candidates at entry level. In addition, the radiologist assigned to the exam committee by the American College of Radiology is invited to attend. In total, 11 subject matter experts participated in the standard setting meeting. Please refer to the following table for specific demographics of this group.



Table 1. Committee Demographics

Rater	Role	Location	Experience	Credentials
A	Sonographer	GA	4-5 years	R.T.(R)(M)(BS)(ARRT)
B	Sonographer	CO	6-10 years	R.T.(R)(M)(BD)(BS)(ARRT)
C	Sonographer	AZ	11-20 years	R.T.(R)(M)(BS)(ARRT)
D	Sonographer	MI	4-5 years	R.T.(R)(M)(BS)(ARRT)
E	Sonographer	NY	6-10 years	R.T.(R)(M)(BS)(ARRT)
F	Sonographer	FL	20+ years	R.T.(R)(M)(BS)(ARRT)
G	Physicist	MI	20+ years	M.S.
H	Sonographer	WA	20+ years	R.T.(R)(M)(BS)(ARRT)
I	Sonographer	CO	20+ years	R.T.(R)(M)(BS)(ARRT)
J	Radiologist*	WI	20+ years	M.D. FACR
K	ARRT Trustee	TX	20+ years	Ph.D., R.T.(R)(QM)(ARRT)

*Committee member J was also serving as an ARRT Trustee at the time of the meeting

Minimally Qualified Candidate

After training regarding the purpose and implications of standard setting, the committee discussed the knowledge, skills, and abilities expected of an entry level breast sonographer with primary focus on the minimum qualifications that should be demonstrated to earn an ARRT credential. This discussion of the “minimally qualified candidate,” who possesses only the knowledge, skills, and abilities required for certification, is important because it allows the committee to come to a common understanding of what is required for the role prior to any data collection activities. Note that “entry level” and “minimally qualified” are not interchangeable terms. Entry level individuals are early in their career with limited clinical experience irrespective of their level of qualification.

Through their discussion, the committee created a list of generic and discipline-specific examples of knowledge, skills, and abilities that are representative of well qualified, minimally qualified, and not yet qualified candidates.

Modified Angoff

The committee performed a modified Angoff activity (Angoff, 1971) using a recently retired exam form. The facilitator provided training to explain the function and intent of the Angoff to the committee and the committee performed a practice activity with a few items to familiarize themselves with the software.

The committee practiced a modified Angoff (Angoff, 1971) procedure with an initial set of twenty items. ARRT staff read each item aloud and committee members independently made judgements for the percentage of minimally qualified candidates that would answer it correctly. The facilitator asked committee members to share their ratings and provide a brief rationale for their judgment. The committee discussed these results with a particular focus on the shared definition of a minimally qualified candidate and clarified their required knowledge, skills, and abilities as necessary.

In the first round of the full activity, the committee reviewed and provided individual judgements for each item on the form. The ratings for each item were averaged across all panelists, and those values were then averaged across all items to determine the minimum percent correct needed to pass the test.

After the first round, the facilitator provided the Angoff-derived cut score for the group as well impact data showing how that score would affect the pass rate of recent candidates. The facilitator then encouraged the committee to discuss their judgments and further clarify their expectations



for the minimally qualified candidate based on real-world experiences with candidates during their coursework or clinical training. The committee reviewed 83 items with high disagreement, possessing a standard deviation greater than 10 as well as an additional 11 items with a standard deviation between 8 and 10.

In round two, committee members reviewed the items again and revised their initial item-level judgments, as desired, based on their discussions during the feedback period.

Table 2. Modified Angoff Results

Percent Correct Cut	Round 1	Round 2
Mean	77.0%	77.6%
Minimum	66.8%	71.2%
Maximum	84.5%	82.4%
Standard Deviation	5.5%	3.4%

Hofstee

The facilitator asked the committee to answer the following four questions to determine the Hofstee recommendation (Hofstee, 1983):

1. What is the lowest acceptable percent correct on the total test you would be comfortable with in order to pass?
2. What is the highest acceptable percent correct on the total test you would be comfortable with in order to pass?
3. What is the minimum percent of test takers that you would be comfortable to fail?
4. What is the maximum percent of test takers that you would be comfortable to fail?

Table 3. Mean Hofstee Responses

Minimum cut score	72%
Maximum cut score	83%
Lowest fail rate	15%
Highest fail rate	37%

Final Discussion

The committee discussed the results of the activities and their overall thoughts regarding the standard setting process and potential impact of a new standard on future pass rates. A summary of the discussion was provided to the ARRT Board of Trustees for review.

ARRT Board of Trustee Review

The ARRT Board of Trustees reviewed the activity results and expressed concern about the number of candidates expected to fail the exam if the standard were increased to the degree suggested. The Board asked ARRT psychometric staff to conduct a second standard setting to provide additional data before committing to a new standard.



Meeting 2 – January 2023

Facilitator:

- Zachary Siegel, Ph.D., Senior Psychometrics Supervisor at ARRT

Committee Composition

ARRT staff selected individuals from the volunteer database with the goal of maximizing diversity in role, geography, and experience in breast sonography. When possible, ARRT will bias the volunteer pool towards individuals early in their career as the exam is designed to assess candidates at entry level. In addition, the radiologist assigned to the exam committee by the American College of Radiology is invited to attend. In total, 12 subject matter experts participated in the standard setting meeting. Please refer to the following table for specific demographics of this group.

Table 4. Committee Demographics

Rater	Role	Location	Experience	Credentials
A	Sonographer	MI	6-10 years	R.T.(R)(M)(BS)(ARRT)
B	Sonographer	MT	4-5 years	R.T.(R)(M)(CT)(BS)(ARRT)
C	Sonographer	NC	1-3 years	MHA, R.T.(R)(M)(BD)(BS)(ARRT)
D	Sonographer	AR	1-3 years	R.T.(R)(M)(CT)(BS)(ARRT)
E	Sonographer	SC	4-5 years	R.T.(R)(M)(BS)(ARRT)
F	Sonographer	HI	1-3 years	R.T.(R)(M)(BS)(ARRT)
G	Sonographer	CO	6-10 years	R.T.(R)(M)(BD)(BS)(ARRT)
H	Sonographer	MN	1-3 years	R.T.(R)(M)(BS)(ARRT)
I	Sonographer	MI	4-5 years	R.T.(R)(M)(BS)(ARRT)
J	Sonographer	NY	6-10 years	R.T.(R)(M)(BS)(ARRT)
K	Sonographer	TN	6-10 years	R.T.(R)(M)(BS)(ARRT)
L	ARRT Trustee	SD	20+ years	R.T.(R)(ARRT)

Minimally Qualified Candidate

After training regarding the purpose and implications of standard setting, the committee discussed the knowledge, skills, and abilities expected of an entry level breast sonographer with primary focus on the minimum qualifications that should be demonstrated to earn an ARRT credential. This discussion of the “minimally qualified candidate,” who possesses only the knowledge, skills, and abilities required for certification, is important because it allows the committee to come to a common understanding of what is required for the role prior to any data collection activities. Note that “entry level” and “minimally qualified” are not interchangeable terms. Entry level individuals are early in their career with limited clinical experience irrespective of their level of qualification.

The committee began their discussion by reviewing a generic list of knowledge, skills, and abilities approved by the ARRT Board of Trustees (Appendix A). The generic list was derived from past standard settings and contains discipline-agnostic expectations for the well qualified, minimally qualified, and not yet qualified candidate. These descriptions were intentionally designed to provide an initial anchor for the committee and facilitate discussion without imposing specific criteria.



After reviewing the generic candidate description, the committee discussed discipline-specific examples of knowledge, skills and abilities that are representative of candidates at each ability level.

Modified Angoff

The committee performed a modified Angoff activity (Angoff, 1971) using a different retired exam form. Although the form differed between meetings, the raw cut score was the same; results presented as a percent correct are therefore comparable between meetings. The facilitator provided training to explain the function and intent of the Angoff to the committee and the committee performed a practice activity with a few items to familiarize themselves with the software.

In the first round of the full activity, committee members read each item on the form and provided their judgment for the percentage of minimally qualified candidates that should answer the item correctly. No additional information was provided during this round.

After the first round, committee members were each provided with feedback regarding their own ratings. Specifically, the facilitator determined each individual's cut score based on the first round before comparing each judgment to the expected percent correct for the individual's cut. The facilitator then provided each committee member with a personalized mix of approximately twenty items that were either rated too high, too low, or close to the expected percent correct for their personal cut score.

The facilitator then provided training for the committee regarding the next round of ratings. Namely, that committee would be able to review the items again with their individual feedback as well as the overall proportion correct for first-time candidates. This training also included a discussion of first-time candidate score distribution and the differing implications of common academic scores (e.g., A, B, C, D, F) and binary pass/fail certification exam results.

Finally, the committee performed a second round of the activity with their first-round judgements provided for them in the response window. Committee members could keep or modify any judgements they desired during this round.

Table 5. Modified Angoff Results

<u>Percent Correct Cut</u>	<u>Round 1</u>	<u>Round 2</u>
Mean	76.6%	79.0%
Minimum	59.4%	64.4%
Maximum	87.2%	87.6%
Standard Deviation	7.0%	6.1%

Hofstee

After the Angoff activity, the facilitator provided training for the Hofstee activity (Hofstee, 1983), directing the committee away from item-level decisions and encouraging them to think more globally about the form they had just reviewed. This training included an explanation of statistical error, impact data, and reasons why the committee may wish to adjust their results higher or lower.



The Hofstee activity consisted of four questions with additional explanations that mirrored the training.

1. What is the lowest acceptable cut score? If a majority of candidates failed the exam, and you were pressured to lower the cut score to permit more candidates to pass, what would be the lowest cut score that you would feel comfortable with? In other words, how low could the cut score be without doing a disservice to the public and the profession (0% - 100%)?
2. What is the highest acceptable cut score? If nearly every candidate passed the exam, and you were pressured to raise the cut score to prevent the exam from being criticized as a trivial standard, what would be the highest cut score that you could live with? In other words, how high could the cut score be without being unfair to candidates and the programs that prepare them (0% - 100%)?
3. What is the lowest acceptable pass rate? Irrespective of the cut score, what is the lowest pass rate that you would be comfortable with (0% - 100%)?
4. What is the highest acceptable pass rate? Irrespective of the cut score, what is the highest pass rate that you would be comfortable with?

Table 6. Mean Hofstee Responses

Minimum cut score	70.6%
Maximum cut score	82.5%
Lowest pass rate	73.5%
Highest pass rate	87.7%

Beuk

The global ratings activity included the Beuk method (Beuk, 1984) alongside the Hofstee method. Committee members answered two questions related to their global expectations for the exam. These questions included additional explanation that mirrored the training.

1. Given everything you know about the content of the exam, and the level of knowledge required for competent performance in a clinical setting, what percentage of the questions should a candidate answer correctly in order to pass the exam (0% - 100%)?
2. Given your knowledge of candidates taking this exam, what percentage of them should pass (0% - 100%)?

Table 7. Beuk Results

Mean cut score	76.3%
SD cut score	5.7%
Mean pass rate	81.7%
SD pass rate	6.3%



Final Discussion

After all data collection was complete, the standard setting committee reviewed the activity results along with an estimated pass rate for potential cut scores within that range. The committee discussed the results, their impressions of the activities, and the cut score they wished to submit to the ARRT Board of Trustees for review. After the discussion, each committee member submitted their final recommended cut score, and the mean of those recommendations was submitted as the overall committee recommendation.

Table 8. Final Recommended Cut Score

Mean	70%
Minimum	68%
Maximum	71%

New Standard and Implementation

The ARRT Board of Trustees reviewed the results and discussed the impact of potential new standards before approving a final standard for the Breast Sonography exam.

The board approved a new standard equivalent to 130 out of 185 items on the exam form used for the second meeting. The new standard will go into effect September 2024 and remain in place until at least 2027 when the next standard setting is scheduled to take place. The new standard is equivalent to seven more correct answers on the exam form used for the second meeting. ARRT staff expect a future pass rate for first-time candidates around 73% based on the impact data provided to both the board and standard setting committee.

References

- Angoff, W.H. (1971). Scales, norms, and equivalent scores. In R.L. Thorndike (Ed.), *Educational Measurement* (2nd ed., pp. 508-600). American Council on Education.
- Beuk, C. H. (1984). A method for reaching a compromise between absolute and relative standards in examinations. *Journal of Educational Measurement*, 21(2), 147-152.
- Hofstee, W. K. (1983). The case for compromise in educational selection and grading. In S.B. Anderson & J.S. Helmick (Eds.), *On educational testing*, (pp.109-127). Jossey-Bass.



Appendix A

Generic Candidate Description Worksheet

At a high level and applicable to all disciplines, this document describes the knowledge, skills, and abilities of clearly passing, clearly failing, and just barely passing candidates. It includes generic examples of items that they would likely get correct or incorrect.

Individual Standard Setting Committees should refine the criteria for the Just Barely Passing candidate specific to their discipline. In doing so, committee members may want to think of entry level people that they know from the workplace or students that they have taught.

Type of Candidate	Typical Knowledge, Skills, and Abilities*	Examples of items that they should answer correctly	Examples of items that they may answer incorrectly
Clearly Passing (way above cut score)	<ul style="list-style-type: none"> • Knows more than what is on the exam • Obtains high quality images on all types of patients • Readily identifies common pathologies • Demonstrates problem-solving (critical thinking) skills • Demonstrates strong oral and written communication skills • Adapts positioning or projections as needed for difficult patients • Anticipates needs of the radiologist • Collaborates with healthcare team to provide optimal patient care 	<ul style="list-style-type: none"> • Fundamental concepts of imaging and patient care • Detailed anatomy (e.g., parts of organs vs whole organs) • Common pathology and physiology • Concrete understanding of how equipment works • Presence and cause of artifacts 	<ul style="list-style-type: none"> • Advanced physics • Unusual anatomic variants • Pharmacology • Advanced calculations/formulas • Difficult item types (multiselect, hot spot, ordered list) • Items with multi-level thinking • May blank out on a simple question (random error)



Type of Candidate	Typical Knowledge, Skills, and Abilities*	Examples of items that they should answer correctly	Examples of items that they may answer incorrectly
Just Barely Passing (just above cut score)	<ul style="list-style-type: none"> • May exhibit test anxiety • Clinical skills greater than or equal to classroom/book knowledge • Possesses limited clinical or work experience (esp. post-primary exams) • Understands basic anatomy and physiology • Can discriminate between normal and abnormal anatomy • Limited knowledge of study indications • Practices quality patient care and safety (not dangerous) • Self-evaluates abilities and knows when to get help • Collaborates with the healthcare team • Unable to explain the “why” (equipment, instrumentation, tube angulation, etc.) 	<ul style="list-style-type: none"> • Patient care • General patient safety • Modality-specific safety guidelines and procedures (radiation, MR, etc.) • Basic image analysis • Definitions • Basic physics • Presence of artifacts • Basic image production items (depending on clinical or lab experience) 	<ul style="list-style-type: none"> • Physics beyond basics • Detailed anatomy (e.g., parts of organs vs. whole organs) • Calculations • Physiology • Cause of artifacts • Correlation with other modalities • Regulatory requirements • QA/QC requirements • Negatively worded items
Failing (below the cut score)	<ul style="list-style-type: none"> • Demonstrates poor test taking skills • Barely passed other exams (post-primary candidates) • Unable to modify protocols from anatomical programming (button pusher) • Unable to adapt positioning or projection based on patient needs or equipment locations • May not understand indications for exams • Weak anatomy and pathology knowledge • Does just enough to meet requirements • Does not know their own limitations • Does not take responsibility for actions • Language barrier 	<ul style="list-style-type: none"> • Basic positioning • Patient care (general concepts) • Consent for exams • Basic modality-specific protection (radiation, MR, etc.) • Basic anatomy only 	<ul style="list-style-type: none"> • Almost anything • Alternate ways to do a study • How the equipment works

*Certain items in the list of knowledge, skills, and abilities are not construct-relevant (e.g., test anxiety, test taking skills, language barrier) but are never-the-less factors that may impact an examinee’s performance on the examination.

