



WHITE PAPER

---

# Certification Exam Validity and Time Optimization

# Table of Contents

EXECUTIVE SUMMARY	3
DETAILED REPORT	4
AAPC CERTIFICATION EXAM CONSTRUCT AND REVIEW	5
BENCHMARKING	8
PSYCHOMETRIC ANALYSIS	10
ADDITIONAL BENEFITS	11
SHORTENING AAPC CERTIFICATION EXAMS	13
RISKS	13
BETA EXAM RESULTS	13
SUMMARY	14
RESOURCES	15
CONTRIBUTORS	15

## Executive Summary

AAPC certification exams are designed to provide certifications for our members to unequivocally demonstrate that the certified individual possesses the critical skills for the position the certification represents. To do this, AAPC must maintain rigorous exams with proven validity and reliability.

AAPC's exam creation and review process ensures the validity of AAPC exams, so that consumers can have confidence that an AAPC certification represents that an individual possesses the critical skills and the expertise the intended position requires. A psychometric review indicates AAPC exams exhibit good to excellent levels of reliability. According to Alan Mead, Ph.D. (psychometrician), "All consumers of AAPC exam scores can have confidence in the pass/fail classifications made using the AAPC exams."

As part of AAPC's exam enhancement and improvement initiative and review of the length of AAPC exams, a benchmark analysis was performed, which compared AAPC exams to medical certification exams and IT certification exams. The benchmark analysis revealed an opportunity to reduce the length of AAPC exams and their administration time. Psychometric analysis confirmed that shorter exams can also exhibit good reliability and content validity.

By shortening AAPC exams, while maintaining the validity, we were able to ensure the same rigorous content as currently provided. Additional benefits of shortening AAPC exams include ensuring adequate time is given to the examinees, reducing obstacles for proctors, and opening an opportunity to provide additional types of questions to further reflect the critical abilities for the intended job.

## Detailed Report

Validity and reliability are common measures indicating the quality of tests.

*Validity: The degree to which accumulated evidence and theory support a specific interpretation of test scores for a given use of a test.*

*Reliability/precision: The degree to which test scores for a group of test takers are consistent over repeated applications of a measurement procedure and hence are inferred to be dependable and consistent for an individual test taker; the degree to which scores are free of random errors of measurement for a given group. (AERA, APA, and NCME, 2014)*

The validity tells us whether the test scores are measuring the right competencies for a particular use of the test. Reliability tells us how consistently the test scores measure something. The goal for any test is high validity and high reliability. If the exam has either poor validity or poor reliability, it is considered a poor test; both measures must be strong.

First, even notable differences in reliability (for a full-length and shortened exam) have small effects on validity (which indexes the utility of the exam scores for a purpose). And second, exams with adequately reliable scores tend to have validity coefficients that are degraded only slightly due to measurement error. Therefore, it seems inappropriate to become overly fixated on meeting a specific threshold for reliability or on the apparent differences for different scenarios. It is more important that all exam scores demonstrate adequate reliability (e.g., above 0.70), because as exam score reliability falls below about 0.70, the criterion-related validity coefficient of the short form will degrade more (and more rapidly as the reliability declines).

Having scores with better reliability is generally advantageous, but not for reasons of validity.

# AAPC Certification Exam Construct and Review

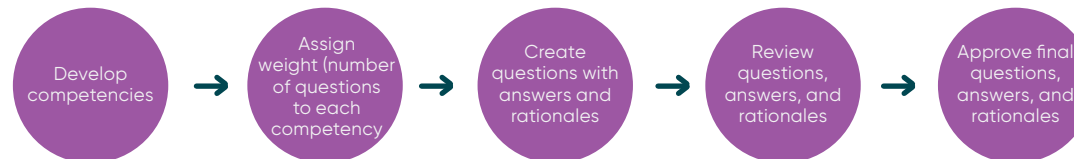
The AAPC process of creating, updating, and reviewing certification exams is done in a way to ensure both validity and reliability.

According to the Standards for Educational and Psychological Testing, there are four phases to tests development:

- 1 Development and evaluation of the test specifications
- 2 Development, tryout, and evaluation of the items
- 3 Assembly and evaluation of new test forms
- 4 Development of the procedures and materials for administration and scoring

## EXAM DEVELOPMENT PROCESS FLOW

### EXAM DEVELOPMENT



### BETA TESTING



### FINAL CERTIFICATION EXAM



Currently, AAPC follows this clear, defined process in the creation of a certification exam blueprint. Once it has been determined that a need for a credential exists, a committee of industry experts is assembled. According to the Standards for Educational and Psychological Testing, “The process of developing education and psychological test should begin with a statement of the purpose(s) of the test, the intended users and uses, the construct or content domain to be measured, and the intended examinee population.” (AERA, APA, and NCME, 2014) This is first on the agenda for the committee. To do this, the committee confirms the need for the certification, identifies the role the credential speaks to, and identifies the job responsibilities for the position. These responsibilities are then translated into domains and competencies required to excel in that position.

The second step in creating a certification exam blueprint is to develop the content specifications. In this step, the committee expands the domains into specific competencies required for the position. The weight of each domain is determined based on the importance of the domain to the position. Competencies are assigned to committee members for question creation.

It is also important for the committee to discuss the format specifications. The type of question (multiple choice, fill-in-the-blank, etc.) is defined based on the demands of the position. Most exam questions are multiple choice with one correct answer and three plausible distractor answer options. A rationale is required on all new questions.

Once created, the questions go through a two-person review (by two people other than the question creator). Each reviewer has a checklist to complete indicating the question and answer are correct and confirming all the requirements have been met. This checklist includes determining if the question tests the competency, confirming the accuracy of the question, validating that the codes are correct for the year of testing, verifying the level of difficulty is appropriate, etc. Questions are then selected for the exam.

For the beta test, applicants are selected, based on their experience, to sit for the beta test and provide feedback. The results from the beta testing are reviewed and modifications are made based on the feedback given. Once the certification exam is finalized, it is placed into production. Through the entire process, checklists follow the exam to ensure the exam is accurate and processes are properly followed.

AAPC certification exams are currently paper-based exams. Examinees have five hours and 40 minutes to complete the exams. The format of each exam is illustrated in Table 1.1.

**TABLE 1.1 AAPC CERTIFICATION EXAM QUESTION FORMATS**

Exam	Exam Name	Number of Questions	Types of Questions
CASCC™	Certified Ambulatory Surgical Center Coder (Specialty)	150	Cases with multiple choice
CANPC™	Certified Anesthesia and Pain Management Coder (Specialty)	150	Cases with multiple choice
CCC®	Certified Cardiology Coder (Specialty)	150	Cases with multiple choice
CCVTC™	Certified Cardiovascular and Thoracic Surgery Coder (Specialty)	150	Cases with multiple choice
CDEO®	Certified Documentation Expert Outpatient (Core)	150	Multiple choice
CEDC®	Certified Emergency Department Coder (Specialty)	150	Cases with multiple choice
CEMC®	Certified Evaluation and Management Coder (Specialty)	150	Cases with multiple choice
CFPC™	Certified Family Practice Coder (Specialty)	150	Cases with multiple choice
CGIC™	Certified Gastroenterology Coder (Specialty)	150	Cases with multiple choice
CGSC™	Certified General Surgery Coder (Specialty)	150	Cases with multiple choice
CHONC™	Certified Hematology and Oncology Coder (Specialty)	150	Cases with multiple choice
CIC®	Certified Inpatient Coder (Core)	70*	Multiple choice Cases with fill-in-the-blank
CIRCC®	Certified Interventional Radiology Cardiovascular Coder (Core)	150	Multiple choice
COBGC™	Certified Obstetrics Gynecology Coder (Specialty)	150	Cases with multiple choice
COPC™	Certified Ophthalmology Coder (Specialty)	150	Cases with multiple choice
COSC™	Certified Orthopaedic Surgery Coder (Specialty)	150	Cases with multiple choice

Exam	Exam Name	Number of Questions	Types of Questions
COC®	Certified Outpatient Coder (Core)	150	Multiple choice
CPEDC™	Certified Pediatric Coder (Specialty)	150	Cases with multiple choice
CPPM®	Certified Physician Practice Manager (Core)	200	Multiple choice
CPB™	Certified Professional Biller (Core)	200	Multiple choice Cases with multiple choice
CPC®	Certified Professional Coder (Core)	150	Multiple choice
CPCD®	Certified Professional Coder in Dermatology Coder (Specialty)	150	Cases with multiple choice
CPCO®	Certified Professional Compliance Officer (Core)	150	Multiple choice
CPMA®	Certified Professional Medical Auditor (Core)	150	Multiple choice Cases with multiple choice
CPMS®	Certified Professional Medical Scribe (Core)	53**	Multiple choice Essay
CRHC™	Certified Rheumatology Coder (Specialty)	150	Cases with multiple choice
CRC®	Certified Risk Adjustment Coder (Core)	150	Multiple choice
CUC™	Certified Urology Coder (Specialty)	150	Cases with multiple choice

\* The CIC® exam is 60 multiple choice questions and 10 cases with multiple fill-in-the-blank questions.

\*\*The CPMS® exam is 50 multiple choice questions and three videos of patient visits for the examinee to scribe.

During exam administration, the examinee has an opportunity to make comments about the exam. All comments are tracked by the Exam Operations Department and responded to by the Exam Content Department, with edits being made to the exams when necessary.

All exams go through a review process to validate the top missed questions and top answered questions regularly. During the update process, external industry experts also review the top missed questions for accuracy and validity. Questions determined to be invalid are omitted from the scoring.

Following these processes has resulted in rigorous, valid, and reliable exams as proven by the psychometric review.

## Benchmarking

Benchmarking was completed to determine how the length of the AAPC exams compares to other exams. In the analysis, data on numbers of items and exam length were randomly sampled from the population of medical- and IT-related exams. Medical-related exams were chosen because these exam programs tend to be conservative and because the AAPC exams have a medical context. The IT-related exams were chosen to show how less conservative exam programs compare to the AAPC exams.

A random sample of 15 exams was selected from each of these results sets. Table 1.2 describes these results. The longest exams sampled were 10 hours long (for two credentials regarding internal medicine and dermatology). A diagnostic radiology credential was about the same (378 minutes). All other exams were shorter than the AAPC exam administration time of 340 minutes. Compared to all sampled exams, the AAPC exams were in the 89<sup>TH</sup> percentile. This means that about 89 percent of all exams have shorter times. Compared to medical exams, the AAPC exams were in the 78<sup>TH</sup> percentile and compared to IT exams, the AAPC exams were in the 97<sup>TH</sup> percentile. These results suggest that the AAPC exams are very long compared to both medical and IT exams. The median exam length for medical exams was 180 minutes (three hours).





The results for numbers of items and minutes per item are harder to interpret, because there is little likelihood that all the items on all exams are of a similar nature. Simple items would require less time, while complex items would require more time. The CPC® exam has 150 items in 340 minutes, or about 2.3 minutes/item, which is near the high end of the range observed for medical exams. Again, this interpretation of the data is difficult since difference kinds of items would require different amounts of time. However, these results suggest that compared to other medical exams, the CPC® items take, relatively speaking, a lot of time to complete.

**TABLE 1.2 DESCRIPTIVE STATISTICS OF TEST LENGTH BASED ON A SAMPLE OF COMPARABLE CERTIFICATION EXAMS**

	Min	Max	Median	Mean	Standard Deviation	Percentile of Full Length AAPC Exam
<b>Healthcare and IT certifications (N=30)</b>						
Time limit (minute)	75	600	150	188.27	124.45	89
Number of items	20	320	120	136.64	78.95	
Time per item (minute)	0.28	7.50	1.47	1.71	1.31	
<b>Healthcare certifications (N=15)</b>						
Time limit (minute)	75	600	180	239.53	157.45	78
Number of items	51	320	180	188.00	72.96	
Time per item (minute)	0.28	2.50	1.03	1.21	0.53	
<b>IT certifications (N=15)</b>						
Time limit (minute)	90	180	150	137.00	30.43	97
Number of items	20	150	80	81.00	35.55	
Time per item (minute)	1.00	7.50	1.88	2.24	1.66	

Note: The number of items was not identified for three of the IT certification exams and two of the healthcare certification exams. They were omitted from calculating the statistics for number of items and time per item.

In conclusion, benchmark samples were randomly sampled from medical and IT exams. Compared to these exams, the AAPC exams were relatively long. The median time for medical exams was three hours and the AAPC exams were longer than about 78 percent of medical exams. When compared to the relatively more progressive domain of IT exams, the AAPC exams were even longer (longer than all sampled exams) and estimated to be the 97<sup>TH</sup> percentile.

## Psychometric Analysis

Cronbach's alpha is an estimate of the psychometric reliability of the exam scores. Psychometric reliability is a proportion (ranging from zero to one: 0.0 to 1.0) of the variability in exam scores due to "true score" variability. A proportion of 1.0 would indicate that all score variability is due to true score variation and 0 percent of the variability is due to error of measurement. This ideal is never realized in practice, but values closer to 1.0 are preferred. While there is no universally accepted standard, reliability below 0.50 suggests very poor-quality measurement (more variability is due to error than true score variation) and values above 0.70 are preferred, with 0.80 being very good and 0.90 being excellent.

The magnitude of coefficient alphas for the AAPC exam scores ranged from .821 to .963 with a mean of 0.899 and median of 0.901 (individual results found in Table 1.3), suggesting that all exams have good to excellent reliability. This means that the exams are precise, as opposed to being influenced by transient error. If transient errors were a significant part of exam scores, an examinee who failed the exam might pass by simply retaking the exam (and many examinees passing the exam would fail if they were retested). This situation does not appear to describe the AAPC exams at all; examinees who pass or fail AAPC exams would be highly likely to be classified identically if they were retested. Thus, all consumers of AAPC exam scores can have confidence in these classifications into master/pass and non-master/fail.

In addition, the reliability of the two forms of the CPC® exam were extremely similar (0.919 and 0.920) and both quite high. This indicates that the quality of the CPC® exam does not vary much across forms and strengthens the argument that it is reasonable to generalize reliability results from a single analyzed form to all comparable current, past, and future forms of an exam.

**Conclusion:** The AAPC exam scores have good to excellent psychometric reliability. Thus, all consumers of AAPC exam scores can have confidence in the pass/fail classifications made using the AAPC exams. An analysis of two forms of the CPC® exam suggests that different AAPC exam forms seem to have very similar results and supports generalizing the results for the analyzed form across all forms for an exam.

**TABLE 1.3 SUMMARY OF ANALYZED AAPC EXAMS**

Exam	Exam Name	Alpha
CPC®	Certified Professional Coder (Core)	.919
		.920
CEMC®	Certified Evaluation and Management Coder (Specialty)	.888
CIC®	Certified Inpatient Coder (Core)	.963
CPPM®	Certified Physician Practice Manager (Core)	.901
COSC™	Certified Orthopaedic Surgery Coder (Specialty)	.901
COC®	Certified Outpatient Coder (Core)	.904
CRC®	Certified Risk Adjustment Coder (Core)	.886
CPB™	Certified Professional Biller (Core)	.933
CPCO®	Certified Professional Compliance Officer (Core)	.821
CIRCC®	Certified Interventional Radiology Cardiovascular Coder (Core)	.941
CPMA®	Certified Professional Medical Auditor (Core)	.859
CDEO®	Certified Documentation Expert Outpatient (Core)	.851

## Additional Benefits

When discussing AAPC exams, it is common to hear that examinees do not have enough time to complete the exam. In 2018, a survey of over 10,000 examinees, 49 percent of the examinees responded that they did not have adequate time to complete the exam. According to the Standards for Educational and Psychological Testing, “test developers should examine the proportion of examinees who complete the entire test ....” To determine this, sign-out times for 200 examinees in 2020 were reviewed to determine those who left with 6-10 minutes left in the exam, and with five minutes left in the exam. The results indicated 23 percent left with 6-10 minutes remaining and 40 percent left in the last five minutes of the exam. Because proctors typically announce when 10 minutes are left in the exam, it is plausible

to presume that examinees leaving within these times finished the exam by randomly filling in the answers to remaining questions. In addition, completed exam grids were reviewed that included comments stating the examinee needed more time on the exam and that they had clearly guessed on an average of 10 questions. The AAPC exams are focus intensive, requiring complex thought processes over an extended period of time. This may result in more of an endurance test than an accurate representation of knowledge.

AAPC local chapter officers may also benefit from a shortened exam. AAPC chapter officers take on many important roles focused on growing and strengthening the chapter they represent and members they serve. AAPC depends on chapter officers to promote the healthcare industry, provide networking opportunities to its members, organize and lead local meetings, and proctor exams. One of the requirements for local chapters is to schedule four exams per year. The officers share responsibility for proctoring the exams or finding volunteers to do so. Most of AAPC's exams are five hours and 40 minutes in duration.

Over the years, it has become increasingly more difficult to secure proctors for the duration of the exam. With a five-hour-and-40-minute testing time frame, proctors usually arrive 30 minutes ahead of the exam start time to prep the area and perform registration tasks. Once the exam has concluded, proctors may remain another 15-30 minutes reorganizing the testing site and performing wrap-up duties. When you add the pre- and post-time it may take to successfully deliver an exam, you may reach or exceed six or seven hours in total duration.

We all are experiencing busier lifestyles. The pressure is on to produce more at our workplace and some find themselves working more than the average 40-hour work week. Because many of our chapter officers have full-time jobs, it takes a strong commitment to also give of their time, especially on weekends when most exams are administered. A study performed by the National Institute for Occupational Safety and Health (NIOSH) reports over 40 percent of study participants admitted to job stress. Couple that with family responsibilities and other personal goals and activities, and it would not be difficult to see how giving six to seven hours of one's time may be difficult at times.

Reducing the time it takes to proctor an exam may actually result in the ability to administer more than one exam per day. This would open additional time slots for examinees and allow proctors to commit less time to a single exam. A time reduction from five hours and 40 minutes to four hours would allow proctors to regain some of their personal time. It may also generate more volunteer interest. Volunteerism is a rewarding act. Many of AAPC members volunteer at organizations such as charities, churches, schools, and other nonprofit organizations (such as AAPC). The number of Americans volunteering has, unfortunately, been on a steady decline over the last two decades (Dietz, 2018). There are many reasons why this has occurred, but one of those reasons is the lack of time many of us have as busy healthcare professionals.

In addition to the benefits mentioned above, reducing the amount of test-taking time also gives chapters the unique opportunity to save or optimize chapter dollars and possibly reallocate funds to other member activities. A recent survey of current chapter officers shows that finding a test site has not been difficult, especially if a chapter officer is employed at a facility that allows them to use the site, such as a hospital or educational institution.

## Shortening AAPC Certification Exams

Shortening exams based on the psychometric data must be done in coordination with subject matter expert (SME) review. In this case, each domain would have the weakest questions removed from each domain at an equal percentage. Based on the number of questions within each domain, this may result in a slight variation of domain percentages. For example, if time were to be cut in half, a domain with five questions would require reducing the domain to two or three questions versus two and a half questions. The questions would need to remain mapped to competencies and sent through SME review to verify all required competencies necessary remained covered in the exam.

Altering the exam to deploy varying question formats on the exam will take more time and effort but may result in increased validity and reliability. For example, the first set of questions would remain multiple choice and the later part of the exam could be cases followed by multiple choice or more real-life situations, such as fill-in-multiple-blanks. This type of shortening of the exams should follow the full new exam creation process.

It is recommended that any process of shortening the exam should be validated with a psychometric review to ensure there is no loss of reliability.

## Risks

The risk of shortening our exams is reducing the reliability or validity. Doing this in conjunction with the psychometric data mitigates the risk of significant loss in reliability. Using the SMEs to ensure all competencies are covered mitigates the risk of loss in validity.

## Beta Exam Results

Following recommendations from the AAPC Exam Board, exam time was reduced to four hours. Using the psychometric data on individual questions, alongside internal AAPC evaluation of individual questions, the item count for the CPC certification exam was reduced to 100 items. Beta exams were taken across our membership base over the period of one month. The results were better than expected by four measurements:

- 1 Reliability: The preliminary psychometric alpha coefficient for the beta exam is higher than anticipated at 0.892. Based on the psychometric review of the full 150-item exam, the projected expectation of a four-hour exam was 0.890.

- ② **Pass Rate:** There was a 1% difference in the pass rate for the 100-item CPC beta exam when compared to the pass rate for the 150-item CPC exam.
- ③ **Beta Exam Survey Results:** The survey results indicate a successful reduction in items and time:
  - 92% agree or strongly agree the questions/cases accurately reflect the proficiency needed in the field for the certification exam.
  - 81% felt the exam was at the appropriate level of difficulty.
  - 79% agree or strongly agree a master's level of proficiency is needed to score well on the exam.
  - 73% felt they were able to answer all questions in the allotted time.
  - 55% felt four hours was the right amount of time.
- ④ **Most Missed Item Review:** The review of the most missed items on the beta exam were consistent with the 150-item exam and the psychometric analysis.

According to Alan Mead, PhD (psychometrician), the results of the psychometric reliability coefficient are good to excellent, and this shows empirically that there is little reason to doubt the projected psychometric results. The reliability for the CPC exam is slightly better than targeted. The pass rate is consistent with the pass rate on a 150-item exam. The most missed questions cover the same top missed competencies typically seen on the 150-item CPC exams. The survey results indicate a successful reaction to the reduction in time.

## Summary

The AAPC certification exam creation and review process has resulted in rigorous, valid, and reliable exams. Psychometric review indicates AAPC's process results in good to excellent levels of reliability on both the 100-item exam and the 150-item exam. Benchmarking data shows that AAPC exams are longer than other similar exams in comparison. Reducing the certification exams to four hours maintains an above average exam time but allows for maintaining reliability standards. Survey results indicate that shortening the exam time and number of items allowed for adequate time to take the exam. Additional benefits of shortening AAPC exams include reducing obstacles for proctors and opening an opportunity to provide additional types of questions to further reflect the critical abilities for the intended job.

In reducing the time allotted for certification exams and the number of items, AAPC will maintain rigorous exams with proven validity and reliability. In addition, AAPC can continue to provide certifications for our members to unequivocally demonstrate that the certified individual has the expertise and meets the skill requirements for the position.

## Resources

American Educational Research Association, American Psychological Association, National Council on Measurement in Education (2014). Standards for educational and psychological testing. Washington, DC: American Educational Research Association.

Dietz, N. (2018, October). "Where are America's Volunteers? A Look at America's Widespread Decline in Volunteering in Cities and States." Retrieved from [https://dogood.umd.edu/sites/default/files/2019-07/Where%20Are%20Americas%20Volunteers\\_Research%20Brief%20\\_Nov%202018.pdf](https://dogood.umd.edu/sites/default/files/2019-07/Where%20Are%20Americas%20Volunteers_Research%20Brief%20_Nov%202018.pdf).

"14 Rules for Writing Multiple-Choice Questions." Brigham Young University, <https://testing.byu.edu/handbooks/14%20Rules%20for%20Writing%20Multiple-Choice%20Questions.pdf>. J. Balogh, PhD (2016). A Practical Guide to Creating Quality Exams. Menlo Park, CA: Intelliphonics, LLC.

NIOSH. (n.d.). "Stress at Work." Retrieved from <https://www.cdc.gov/niosh/docs/99-101/pdfs/99-101.pdf?id=10.26616/NIOSH PUB99101>.

"Validating Your Certification Exam." Association of Boards of Certification, [www.abccert.org/pdf\\_docs/ValidatingYourCertificationExam.pdf](http://www.abccert.org/pdf_docs/ValidatingYourCertificationExam.pdf).

Zimmaro, Dawn M. "Writing Good Multiple-Choice Exams," <https://facultyinnovate.utexas.edu/sites/default/files/writing-good-multiple-choice-exams-fic-120116.pdf>.

## Contributors

**Katherine Abel**, CPC, CPB, CPMA, CPPM, CPC-I, AAPC Fellow  
Sr. Director of Content Development | AAPC

**Nicole Benjamin**, CPC, CPC-I, CEDC  
Coding Education Specialist | AAPC

**Leesa A. Israel**, BA, CPC, CUC, CEMC, CPPM, CPC-I, CMBS  
Head of Publishing, Editorial & Technology | AAPC

**Jaci Johnson Kipreos**, CPC, CPMA, CDEO, CEMC, COC, CPC-I  
President | Practice Integrity, LLC

**Alan D. Mead**, Ph.D.  
President | Talent Algorithms Inc

**Shantanu Phadnis**  
VP Content and Curriculum | AAPC

**Dave Walker**  
Solutions Engineer, Technical Development | AAPC

**Lee Williams**, MBA, CPC, CHONC, CEMC, CPCO, CRC, RHIA, CCS, CCDS  
Director of Education | AAPC

**Wendy Willes**  
Exam Department Head | AAPC

Register for your next exam at [www.aapc.com/certification/](http://www.aapc.com/certification/)