

NORTH AMERICAN REGISTRY OF MIDWIVES

EXAMINATION PROGRAM TECHNICAL REPORT 2021

May 2021

CONTENTS

	<u>Page</u>
Introduction	4
Section I: Examination Administration	5
Section II: General Test Results	6
Section III: Scaled Scores	13
Section IV: The Skills Assessment Examination	14
Section V: Test Development	15

Tables

	<u>Page</u>
Table 1: Examination Statistics	7
Table 2: Pass/Fail Frequency Distribution	8
Table 3: Frequency Distribution of Form R Raw Scores	9
Table 4: Frequency Distribution of Form S Raw Scores	10
Table 5 Test Statistics: Form R and Form S – Section 1	11
Table 6 Test Statistics: Form R and Form S - Section 2	11
Table 7 Test Statistics: Form R and Form S - Section 3	11
Table 8 Test Statistics: Form R and Form S - Section 4	12
Table 9 Test Statistics: Form R and Form S - Section 5	12
Table 10 Test Statistics: Form R and Form S - Section 6	12
Table 11 Test Statistics: Form R and Form S - Section 7	13

Introduction

This report provides an overview of examination program activities and the statistical properties of Forms R and S of the Certified Professional Midwife (CPM) Examination, administered by the *North American Registry of Midwives* (NARM). Successful completion of the CPM examination must be attained before the title of Certified Professional Midwife is conferred.

The forms of the examination discussed in this report were administered via remote on-line proctoring and via computer in a number of locations at different times throughout the United States during 2020. The data from these administrations were combined and served as the basis for the statistics appearing in this report.

SECTION I: EXAMINATION ADMINISTRATION

The examination was administered via a computer-based testing (CBT) format, both in-person and remote proctored. Two forms of the examination were administered; Form R and Form S.

	Examination Forms	
	R	S
Number of Examinees	134	151

This pool of candidates taking Forms R and S were used as the basis of the analyses appearing in this report.

SECTION II: GENERAL TEST RESULTS

Each test form consists of a total of 300 questions contained in two separate parts of 150 questions each. The examinations were administered in two 3-hour sessions, morning and afternoon. The mean, standard deviation, reliability coefficient and standard error of measurement for Form R and Form S are based on this total of 300 items (see Table 1). The raw score mean is the average number of items answered correctly by the group of examinees. For example, the table shows that the mean or average raw score for Form R was 252.46.

The Kuder-Richardson-20 (or KR-20) reliability reflects the degree of consistency in the test scores.

The Standard Error of Measurement is interpreted as a standard deviation of the errors of measurement for the test, and is directly influenced by both the size of the standard deviation and the degree of unreliability of the test. For Form R the standard error of measurement is equal to 5.71. The greater the standard error of measurement, the more the score reflects chance factors.

Table 2 summarizes the pass/fail statistics for Form R and Form S in terms of both the number of candidates and the percentage of candidates. For example, 99 of the 134 candidates (or 73.9%) who took Form R passed the exam.

Table 3 and Table 4 present the frequency distributions of the raw scores for Form R and Form S, along with univariate statistics including the mean, median, mode, standard deviation, skewness and kurtosis statistic. Skewness measures the extent to which the scores are symmetric about the mean. The value obtained for Form R, for example (-0.78761) indicates that the distribution of scores is slightly skewed to the left (i.e., "negatively skewed"). Kurtosis measures the flatness of a distribution or the heaviness of its tails. The standard is the normal distribution with a value of 0. Distributions with short tails and few extreme scores have negative kurtosis. The positive kurtosis indicated for Form R points to a distribution with a larger number of scores more distant from the mean. The frequency distributions in Table 3 (Form R) and Table 4 (Form S) include the number, percentage, and cumulative percentage of candidates who obtained each raw score. For instance, in Table 3, we can see that 3 examinees (or 2.1% of the total pool of 134 candidates) obtained a raw score of 238 on Form R of the examination, and that approximately 21.1% of the candidates achieved a score of 238 or lower.

Each form of the test consists of seven sections. The following table indicates the number of items in each of the seven sections of Form R and Form S:

Section	Number of Items	
	Form R	Form S
1	13	13
2	16	16
3	28	28
4	79	79
5	107	107
6	40	40
7	17	17
Total	300	

Tables 5 through 11 present test statistics for each of the seven sections including the raw score mean and median, standard deviation and KR-20 reliability coefficient.

Table 1: Examination Statistics

	Form R	Form S
Number of Scored Items	300	300
Total Number of Examinees	134	151
Raw Score Mean	252.46	252.98
Raw Score Standard Deviation	21.60	21.67
KR-20 Reliability Coefficient	0.93	0.93
Standard Error of Measurement	5.71	5.73
Raw Passing Score	243	243
Percent Passing Score	81.00	81.00
Decision Consistency (Livingston)	0.94	0.94

Table 2: Pass/Fail Frequency Distribution

	Form R	Form S
Pass	99 (73.9%)	112 (74.2%)
Fail	35 (26.1%)	39 (25.8%)
Total	134 (100.0%)	151 (100.0%)

Table 3: Frequency Distribution of Form R Raw Scores

Raw Score	Frequency	Cumulative Frequency	Percent	Cumulative Percent
190	1	1	0.7	0.7
193	1	2	0.7	1.4
207	1	3	0.7	2.1
209	2	5	1.5	3.6
214	2	7	1.5	5.1
216	2	9	1.5	6.6
219	2	11	1.5	8.1
221	1	12	0.7	8.8
223	1	13	0.7	9.5
226	2	15	1.5	11.0
228	4	19	2.9	13.9
231	2	21	1.5	15.4
233	1	22	1.5	16.9
235	3	25	2.1	19.0
238	3	28	2.1	21.1
240	4	32	2.9	24.0
242	3	35	2.1	26.1
245	8	43	6.0	32.1
247	4	47	3.0	35.1
250	14	61	10.4	45.5
252	5	66	3.7	49.2
254	4	70	3.0	52.2
257	6	76	4.5	56.7
259	8	84	6.0	62.7
261	4	88	3.0	65.7
264	8	96	6.0	71.7
266	7	103	5.2	76.9
269	9	112	6.7	83.6
271	2	114	1.5	85.1
273	10	124	7.5	92.6
276	4	128	3.0	95.6
278	2	130	1.5	97.1
280	2	132	1.5	98.6
283	1	133	0.7	99.3
285	1	134	0.7	100.0

Sample Size: 134**Minimum:** 190**Mean:** 252.46**Mode:** 250**Standard Deviation:** 21.60**Skewness:** -0.78761**Maximum:** 285**Median:** 253.0**Kurtosis:** 0.420076

Table 4: Frequency Distribution of Form S Raw Scores

Raw Score	Frequency	Cumulative Frequency	Percent	Cumulative Percent
174	1	1	0.7	0.7
197	1	2	0.7	1.4
202	1	3	0.7	2.1
205	1	4	0.7	2.8
212	2	6	1.3	4.1
219	1	7	0.7	4.8
221	1	8	0.7	5.5
224	2	10	1.3	6.8
226	2	12	1.3	8.1
228	4	16	2.6	10.7
231	8	24	5.3	16.0
233	2	26	1.3	17.3
235	5	31	3.3	20.6
238	4	35	2.6	23.2
240	4	39	2.6	25.8
243	8	47	5.3	31.1
245	4	51	2.6	33.7
247	8	59	5.3	39
250	7	66	4.6	43.6
252	8	74	5.3	48.9
254	8	82	5.3	54.2
257	3	85	2.0	56.2
259	6	91	4.0	60.2
262	5	96	3.3	63.5
264	7	103	4.6	68.1
266	7	110	4.6	72.7
269	6	116	4.0	76.7
271	6	122	4.0	80.7
273	9	131	6.1	86.8
276	7	138	4.6	91.4
278	4	142	2.6	94.0
281	3	145	2.0	96.0
283	4	149	2.6	98.6
285	1	150	0.7	99.3
290	1	151	0.7	100.0

Sample Size: 151**Minimum:** 174**Maximum:** 290**Mean:** 252.98**Median:** 252.5**Mode:** 273**Standard Deviation:** 21.67**Skewness:** -0.46465**Kurtosis:** -0.17855

Table 5: Test Statistics: Form R and Form S – Section 1

Statistic	Form R	Form S
Number of Tests Scored	134	151
Number of Scored Items	13	13
Mean Score	11.40	11.50
Median Score	11.00	12.00
Standard Deviation	1.11	1.25
K-R 20 Reliability Coefficient	0.40	0.45

Table 6: Test Statistics: Form R and Form S – Section 2

Statistic	Form R	Form S
Number of Tests Scored	134	151
Number of Scored Items	16	16
Mean Score	13.13	13.34
Median Score	13.00	14.00
Standard Deviation	2.01	1.63
K-R 20 Reliability Coefficient	0.56	0.48

Table 7: Test Statistics: Form R and Form S – Section 3

Statistic	Form R	Form S
Number of Tests Scored	134	151
Number of Scored Items	28	28
Mean Score	23.26	23.43
Median Score	24.00	24.00
Standard Deviation	2.86	2.90
K-R 20 Reliability Coefficient	0.61	0.61

Table 8: Test Statistics: Form R and Form S – Section 4

Statistic	Form R	Form S
Number of Tests Scored	134	151
Number of Scored Items	79	79
Mean Score	65.42	64.87
Median Score	67.00	66.00
Standard Deviation	6.50	6.66
K-R 20 Reliability Coefficient	0.79	0.79

Table 9: Test Statistics: Form R and Form S – Section 5

Statistic	Form R	Form S
Number of Tests Scored	134	151
Number of Scored Items	107	107
Mean Score	92.17	92.11
Median Score	94.00	93.00
Standard Deviation	7.94	8.19
K-R 20 Reliability Coefficient	0.82	0.83

Table 10: Test Statistics: Form R and Form S – Section 6

Statistic	Form R	Form S
Number of Tests Scored	134	151
Number of Scored Items	40	40
Mean Score	33.60	33.87
Median Score	34.00	34.00
Standard Deviation	3.53	3.54
K-R 20 Reliability Coefficient	0.64	0.66

Table 11: Test Statistics: Form R and Form S – Section 7

Statistic	Form R	Form S
Number of Tests Scored	134	151
Number of Scored Items	17	17
Mean Score	13.48	13.86
Median Score	14.00	14.00
Standard Deviation	1.88	1.79
K-R 20 Reliability Coefficient	0.39	0.37

SECTION III: SCALED SCORES

The following conversion formula was used for determining the equivalent scaled scores for each raw score. The scaled scores are expressed over a range of 0 to 99, with 75 as passing.

Form R and Form S:

$$\text{Scaled Score} = .4210526 \times \text{Raw Score} - 27.31578$$

(In the unlikely event of a scaled score below zero, it will be reported as zero.)

SECTION V: TEST DEVELOPMENT

In 1986, Midwives Alliance of North America established the Interim Registry Board (IRB) to develop a test that would measure midwifery knowledge based on the MANA Core Competencies being developed by the MANA Education Committee. In 1991, the first test was administered to groups of midwives across the United States. This “trial” exam was revised under the guidance of a testing consultant, and by November of 1991, it was officially administered as the North American Registry of Midwives Written Examination. With yearly revisions, the Registry Examination continued to be administered and those who passed were listed on the “Registry.”

In 1992, The Interim Registry Board was reformed as a non-profit corporation separate from MANA. The new organization was named the North American Registry of Midwives, more commonly referred to as NARM.

The first CPM credential was issued in November of 1994. This marked the end of the NARM Registry process (list of those who passed the exam) and the beginning of the NARM Certification Process (verification of education and experience, and passing the exam).

The 1995 NARM Job Analysis was undertaken by NARM in conjunction with the National Assessment Institute.

In 1996, NARM signed a full contract with Schroeder Measurement Technologies that included test development for Form D, test administration, and the administration of the portfolio application process. The earlier versions of the exam (Forms A-C) had evolved from a combination of multiple choice and essay, to a total multiple choice format of 350 questions.

In 1997, NARM contracted with Dr. Gerald Rosen of National Measurement and Evaluation (NME) to function as the testing company. National Measurement and Evaluation worked with NARM to continue both the Job Analysis in 2000-2001, and again in 2008-2009, and to regularly update the exam forms. These Job Analysis surveys were done by mailing survey forms and scanning the returned forms.

Dr Rosen has continued as NARM’s primary psychometrician, though the 2015-2016 Job Analysis Survey was done electronically under the direction of Ellen Julian and David Paulson with Inteleos Psychometric Services.

The NARM Test Development process functions in a 3-4 year cycle, with Item Writing workshops being hosted around the country for two years, a third year for review of new items, and a fourth year for implementation of new forms. Items are written during two-day workshops with several teams of three Subject Matter Experts writing approximately 20 new items per workshop. These item writers also serve as review teams for items written by previous teams. Each proposed item is reviewed by at least two item writer teams, a third time by the professional members of the board of directors, and a fourth time by the psychometrician.

Approved new items are added the Itembank, and two new forms of the exam are developed using the new items as core items on both forms and other items matched for topic, domain, and difficulty level to be on equivalent exam forms. The two new forms are evaluated by Subject Matter Experts during an Angoff Cut Score workshop led by Dr Rosen. The board reviews the recommendations and has the option to lower the cut score by one standard deviation.

Two new forms of the exam will be offered in September, 2021. These exams are the result of Item Writing workshops in 2018-2019. Ongoing development was delayed due to Covid in 2020, so the reviews and cut scores were done virtually in 2021. Sixty-five CPMs attended one of the five Item Writing workshops during this two-year period, and also served to review items written by other item writers.