

Category 3 Reporting Guidelines for P4P Quality of Life and Behavioral Health Tools with a Pretest/Posttest Survey Design

The following guidance is for Category 3 Quality of Life (OD-10) and Behavioral Health tools (OD-11) measures that are classified as Pay for Performance (P4P). This methodology is specific to tools that are used to longitudinally track individual progress, and is not intended for use with patient satisfaction measures (OD-6), or the CDC Health Related Quality of Life measure (IT-10.1.h). For reporting guidelines on OD-6 or IT-10.1.h, see the relevant Category 3 Compendium documents.

This guidance explains options for baseline reporting and determining DY4 and DY5 improvement targets, as well as guidance for determining the optional Average Patient Population Score Boundary.

This guidance applies to the following measures only:

IT-10.1.a.i	AQoL-4D	IT-10.1.b.ii	RAND SF-12	IT-10.3.e	PAID
IT-10.1.a.ii	AQoL-6D	IT-10.1.b.iii	RAND SF-36	IT-11.26.e.i	PHQ-9
IT-10.1.a.iii	AQoL-7D	IT-10.1.f	FACT-G	IT-11.26.e.ii	PHQ-15
IT-10.1.a.iv	AQoL-8D	IT-10.2.b	IADLs Scale	IT-11.26.e.iii	PHQ-SADS
IT-10.1.a.v	PedsQL	IT-10.3.d	BDI-2	IT-11.26.e.iv	PHQ-4

Providers will determine their baseline and DY4 and DY5 achievement levels using one of the following scenarios. Providers will report their selected scenario as well as their retest period as part of their survey administration description requested as supporting documentation for baseline reporting.

Providers may not switch between scenarios in subsequent measurement years.

Scenario 1: Baseline includes pre and posttest scores, improvement in change score

Baseline	average posttest score and average pretest score reported to determine change score
DY4 improvement target	Baseline average posttest + 5% of change score
DY5 improvement target	Baseline average posttest + 10% of change score

Scenario 2: Baseline includes pretest scores only, improvement over baseline score

Baseline	average pretest score
DY4 improvement target	Baseline average posttest + 5% of range of possible survey scores
DY5 improvement target	Baseline average posttest + 10% of range of possible survey scores

Scenario 3: No pre/post-testing methodology, improvement over baseline score

Baseline	average survey score
DY4 improvement target	Baseline average survey score + 5% of range of possible survey scores
DY5 improvement target	Baseline average survey score + 10% of range of possible survey scores

Helpful Definitions:

Pretest: The first survey issued to an individual, before or at the beginning of DSRIP intervention services

Posttest/Retest: Survey given after DSRIP services have begun or completed, to demonstrate improvements in the individual/target population.

Retest Period: The time between issuing the pretest and posttest or subsequent testing (i.e. testing frequency).

Average Most Recent Score: The most recent posttest score. As some providers may issue more than one posttest, average most recent score is used to indicate that the most recent posttest score should be used for individuals who have received more than one posttest.

Directionality: All tools are classified as having either negative or positive directionality. Negative directionality is when lower scores indicate better outcomes, and providers will need to lower their average reported score in DY4/DY5 to meet improvement target goals. Positive directionality is when higher survey scores indicate better outcomes, and providers will need to increase average reported score in DY4/DY5 in order to meet improvement target goals.

P4P Scenario 1: Baseline includes pre and posttest scores

Reporting Baseline for P4P Scenario 1

For baseline reporting, providers will report the following two scores:

1. The average pretest score of all individuals who have received at least two surveys since the beginning of DY1, with the most recent survey completed during the baseline measurement period.
2. The average most recent score of all individuals who received at least two surveys during the baseline measurement period.

If an individual received a pretest only, their score should be excluded from the pretest baseline score as these individuals will not be represented in the average most recent score. In determining baseline, providers should follow measure specifications per the survey developer for retest periods. Where no retest periods are specified, providers should assume a reasonable retesting period based on clinical standards and guidelines. Providers will report retest periods to HHSC during baseline reporting.

Reporting DY4 & DY5 for P4P Scenario 1

In DY4, providers will report the average most recent score of individuals who received at least two surveys since the beginning of the baseline measurement period, and whose most recent survey was completed during the DY4 measurement period.

In DY5, providers will report the average most recent score of individuals who received at least two surveys since the beginning of the baseline measurement period and whose most recent survey was completed during the DY5 measurement period.

For individuals who receive more than one posttest survey during the course of the demonstration, only the most recent score should be reported in a given measurement year.

Finding DY4 & DY5 Improvement Targets for P4P Scenario 1

In DY4 and DY5, the improvement targets will be determined by the difference between the DY3 average pretest score, and DY3 average most recent score.

In DY4, providers will improve by 5% over DY3 baseline the difference between the DY3 reported average pretest score and the average most recent score of individuals who received at least two tests since the start of the baseline measurement period and whose most recent test was completed during the DY4 measurement period.

DY4 improvement target for positive directionality:

DY3 average pretest score + 1.05(DY3 average most recent score - DY3 average pretest score)

DY5 improvement target for negative directionality:

DY3 average pretest score - 1.05(DY3 average pretest score - DY3 average most recent score)

In DY5, providers will improve by 10% over DY3 baseline the difference between the DY3 reported average pretest score and the average most recent score among individuals who received at least two tests since the start of the baseline measurement period and whose most recent test was completed during the DY5 measurement period.

DY5 improvement target for positive directionality:
DY3 average pretest score + 1.10(DY3 average most recent score - DY3 average pretest score)

DY5 improvement target for negative directionality:
DY3 average pretest score - 1.10(DY3 average pretest score-DY3 average most recent score)

See Examples D & E at the end of this document for an example of Scenario 1 Baseline and DY4 and DY5 reporting.

P4P Scenario 2: Baseline includes pretest scores only

Reporting Baseline for P4P Scenario 2

For baseline reporting, providers will report the average pretest score of individuals who received a pretest survey during the measurement year.

Providers should follow measure specifications per the survey developer for retest periods. Where no retest periods are specified, providers should assume a reasonable retesting period based on clinical standards and guidelines. Providers will report retest periods to HHSC during baseline reporting. If provider issues posttests during baseline measurement, those scores should not be included in the baseline calculation.

Reporting DY4 & DY5 for P4P Scenario 2

In DY4, providers will report the average most recent score of individuals who received at least two surveys since the beginning of the baseline measurement period and whose most recent test was completed during the DY4 measurement period.

In DY5, providers will report the average most recent score of individuals who received at least two surveys since the beginning of the baseline measurement period and whose most recent test was completed during the DY5 measurement period.

For individuals who receive two or more surveys during the course of the demonstration, only the most recent score should be reported in a given measurement year.

Finding DY4 & DY5 Improvement Targets for P4P Scenario 2

In DY4: the improvement target will be an improvement over baseline equal to 5% of the difference between the maximum and minimum point values of the selected survey tool.

DY4 improvement target for positive directionality:
DY3 average pretest score + .05*(maximum possible survey score - minimum possible survey score)

DY4 improvement target for negative directionality:
DY3 average pretest score - .05*(maximum possible survey score - minimum possible survey score)

In DY5: the improvement target will be an improvement over baseline equal to 10% of the difference between the maximum and minimum point values of the selected survey tool.

DY5 improvement target for positive directionality:

DY3 average pretest score + .10*(maximum possible survey score - minimum possible survey score)

DY5 improvement target for negative directionality:

DY3 average pretest score - .10*(maximum possible survey score - minimum possible survey score)

Example A:

Provider has selected a tool with a minimum score of 0 and a maximum score of 60 with positive directionality where higher numbers represent better outcomes. In DY3, provider finds that their average pretest score is 42.

DY4 Improvement Target = 42 + .05(60-0) = 45*

DY5 Improvement Target = 42 + .10(60-0) = 48*

In this example, the providers DY4 improvement target is an average most recent score of 45. The providers DY5 improvement target is an average most recent score of 48.

P4P Scenario 3: No pre/posttest methodology

Reporting Baseline for P4P Scenario 3

For baseline reporting, providers will report the average score of all surveys completed during the baseline measurement period.

Reporting DY4 & DY5 for P4P Scenario 3

In DY4, providers will report the average score of all surveys completed during the DY4 measurement year.

In DY5, providers will report the average score of all surveys completed during the DY5 measurement year.

Finding DY4 & DY5 Improvement Targets for P4P Scenario 3

In DY4 the improvement target will be an improvement equal to 5% of the difference between the maximum and minimum point values of the selected survey tool.

DY4 improvement target for positive directionality:

DY3 average score + .05*(maximum possible survey score - minimum possible survey score)

DY4 improvement target for negative directionality:

DY3 average score - .05*(maximum possible survey score - minimum possible survey score)

In DY5 the improvement target will be an improvement equal to 10% of the difference between the maximum and minimum point values of the selected survey tool.

DY5 improvement target for positive directionality:

DY3 average score + .10*(maximum possible survey score - minimum possible survey score)

DY5 improvement target for negative directionality:
DY3 average score - .10*(maximum possible survey score - minimum possible survey score)

Example B:

Provider has selected a tool with a minimum score of 15 and a maximum score of 85 with positive directionality where higher numbers represent better outcomes. In DY3, provider finds that their average score is 51.

$$\text{DY4 Improvement Target} = 51 + .05*(85-15) = \mathbf{54.5}$$

$$\text{DY5 Improvement Target} = 51 + .10*(85-15) = \mathbf{58}$$

In this example, the providers DY4 improvement target is an average score of 54.5. The providers DY5 improvement target is an average score of 58.

Which P4P Scenario Should Providers Use?

Scenario 1: Baseline includes pre and posttest scores, improvement in change score

Scenario 1 requires that providers are able to report both pre and posttest scores in their baseline measurement period, and allows providers to set their DY4 and DY5 improvement targets based on their initially observed improvement levels. Providers may consider this scenario when the improvement targets in Scenario 2 and Scenario 3 aren't meaningful or realistic (if the 5% or 10% improvement targets are too high or too low) for a given project/population. This scenario may be a good fit for providers that will be intervening on multiple and/or similar intervention cohorts and where the impact of the intervention on scores has not yet been established. Providers can meet their DY4 & DY5 improvement targets by increasing the amount of improvement seen in individuals between pre-and posttest and/or increasing the percentage of individuals who show improvement.

Scenario 2: Baseline includes pretest scores only, improvement over baseline score

Scenario 2 should be considered for projects where only pretest scores are available in DY3 because of long retest periods, or the provider is following a single or substantially similar cohort across multiple years. Improvement targets are fixed to the total available points in a selected survey, regardless of initial baseline performance as DY4 and DY5 goals are set as a 5% and 10% improvement of the total survey points over the average pretest score.

Scenario 3: No pre/post-testing methodology, improvement over baseline score

Scenario 3 should be considered for projects where survey recipients will not be retested, or if the provider does not have the capacity to stratify data as pre and posttest at the individual level.

Category 3 Reporting Guidelines for P4R Quality of Life and Behavioral Health Tools with a Pretest/Posttest Survey Design

The following guidance is for DSRIP Category 3 Quality of Life (OD-10) and Behavioral Health (OD-11) Tool 3 measures that are classified as Pay for Reporting (P4R). This methodology is specific to tools that are used to track individual progress over time, and is not intended for use with a patient satisfaction measures (OD-6), or the CDC Health Related Quality of Life measure (IT-10.1.h). For reporting guidelines on OD-6 or IT-10.1.h, please reference relevant compendium documents. This guidance explains options for baseline reporting and DY4/ DY5 reporting, and applies to the following measures:

IT-10.1.c	Q-LES-Q
IT-10.1.i.ii	CHQ-PF28
IT-10.2.a	SIS
IT-10.4.a	DP-3
IT-10.4.b	VABS II

IT-10.5	Bayley-III
IT-11.24	GAD-7
IT-11.25	DLA-20
IT-11.26.b	ABC
IT-11.26.c	ANSA

IT-11.26.d	CANS-MH
IT-11.26.e.v	EPD

Providers will report their baseline and DY4 and DY5 achievement levels using one of the following scenarios, determined by the scores reported at baseline reporting. Providers will report their selected scenario, as well as their retest period as part of their survey administration description supporting documentation required for baseline reporting.

Providers may not switch between scenarios in subsequent measurement years.

Scenario 1: Baseline includes pre and posttest scores

In DY4 and DY5, provider reports average posttest scores

Scenario 2: Baseline includes pretest scores only

In DY4 and DY5, provider reports average posttest scores

Scenario 3: No pre/post-testing methodology

In baseline period, DY4 and DY5, provider reports average survey score in each measurement year

P4R Scenario 1: Baseline includes pre and posttest scores

Reporting Baseline for P4R Scenario 1

For baseline reporting, providers will report the following two scores:

1. The average pretest score of all individuals who have received at least two surveys since the beginning of DY1, with the most recent survey completed during the baseline measurement period.
2. The average most recent score of all individuals who received at least two surveys during the baseline measurement period.

If an individual received a pretest only, their score should not be counted in the pretest baseline score calculation. In determining baseline, providers should follow measure specifications for retest periods. Where no retest periods are indicated, providers should assume a reasonable retesting period based on clinical relevance. Providers will report retest periods to HHSC during baseline reporting.

Reporting DY4 & DY5 for P4R Scenario 1

In DY4, providers will report the average most recent score of individuals who received at least two surveys since the beginning of the baseline measurement period and whose most recent survey was completed during the DY4 measurement year.

In DY5, providers will report the average most recent score of individuals who received at least two surveys since the beginning of the baseline measurement period and whose most recent survey was completed during the DY5 measurement year.

For individuals who receive more two or more surveys during the course of the demonstration, only the most recent score should be reported in a given measurement year.

P4R Scenario 2: Baseline includes pretest scores only

Reporting Baseline for P4R Scenario 2

For baseline reporting, providers will report the average pretest score of individuals who received a pretest survey during the measurement year.

Providers should follow measure specific specifications for retest periods. Where no retest periods are indicated, providers should assume a reasonable period between pretest and retest based on clinical relevance. Providers will report retest periods to HHSC during baseline reporting. If provider issues posttests during baseline measurement, posttest scores should not be included in the baseline.

Reporting DY4 & DY5 for P4R Scenario 2

In DY4, providers will report the average most recent score of individuals who received at least two surveys since the beginning of the baseline measurement period and whose most recent test was completed during the DY4 measurement period.

In DY5, providers will report the average most recent score of individuals who received at least two surveys since the beginning of the baseline measurement period and whose most recent test was completed during the DY5 measurement period.

For individuals who receive two or more surveys during the course of the demonstration, only the most recent score should be reported in a given measurement period.

P4R Scenario 3: No pre/posttest methodology

Reporting Baseline for P4R Scenario 3

For baseline reporting, providers will report the average score of all surveys completed during the baseline measurement period.

Reporting DY4 & DY5 for P4PR Scenario 3

In DY4, providers will report the average score of all surveys completed during the DY4 measurement year.

In DY5, providers will report the average score of all surveys completed during the DY5 measurement year.

Pretest Score Boundary for Cat 3 Quality of Life and Behavioral Health Tools with a Pretest/Posttest Survey Design

During baseline measurement period, providers using P4P or P4R Scenario 1 or 2 may establish a pretest score boundary to define their average population. The score boundary allows providers to report the most recent score of individuals whose pretest score falls within a specified range, excluding individuals with very high and low scores, and allowing providers to report progress within similar populations throughout measurement years.

To determine the pretest score boundary, providers will issue the tool/survey to their full population of individuals receiving pretest assessment, and use the full set of pretest scores to determine their population mean and standard distribution. [Finding the mean and standard distribution in excel](#)

Pretest Score Boundary:

Minimum Score = DY3 All Scores Pretest Mean - (1 or 2)*DY3 Standard Deviation

Maximum Score = DY3 All Scores Pretest Mean + (1 or 2)*DY3 Standard Deviation

Providers may elect to use a score boundary of one or two standard deviations depending on population size, as providers using a score boundary are still required to report the minimum sample size required for Category 3 measurement periods. The score boundary will not change in subsequent reporting years, even if a provider's mean and standard distribution change.

P4P & P4R Scenario 1:

- Baseline: Providers will report the average pretest and the average most recent score of individuals whose pretest score was within the average patient population score boundary.
- DY4 & DY5: Providers will report the average most recent score of individuals whose pretest score was within the score boundary.

P4P & P4R Scenario 2:

- Baseline: Providers will report the average pretest score of individuals whose pretest score was within the average patient population score boundary.
- DY4 & DY5: Providers will report the average most recent score of individuals whose pretest score was within the average patient population score boundary.

Providers should continue to posttest individuals whose pretest scores are outside the pretest score boundary if it is clinically relevant to their intervention, but the most recent scores of individuals whose pretest scores are out of range should not be used for calculating Category 3 reporting values.

Providers with large patient populations may elect to conduct a random sample of their population to determine their pretest score boundary. Providers using a pretest score boundary should follow compendium guidelines for denominator size.

Providers with small populations must still report a minimum of 30 cases per measure during a 12 month measurement period or 15 cases for a 6 month measurement period. Provider should only use a pretest score boundary if doing so results in a population large enough to report the minimum number of cases required by the compendium guidelines.

Example C: Determining Pretest Score Boundary for Scenarios 1 & 2

	DY3
Case #	Pretest
1	32
2	23
3	47
4	10
5	44
6	11
7	52
8	7
9	12
10	31
11	34
12	17
13	12
14	33
15	49
16	24
17	43
18	15
19	24
20	31
21	14
22	58
23	47
24	9
25	13
26	39
27	14
28	39
29	10
30	11
31	18
32	50
33	31
34	48
35	35
MEAN	29.5

In the example to the left, a provider selected a quality of life measure with a minimum score of 1 and a maximum score of 60, where higher numbers reflect a better quality of life. After issuing the full population a pretest during the DY3 baseline measurement period, the provider determined the following:

DY3 Mean Pretest Score (All Scores): 29.5

DY3 Standard Deviation (SD): 17.6

Min Score = DY3 Pretest Mean - DY3 SD = 29.5 - 17.6 = 11.9

Max Score = DY3 Pretest Mean + DY3 SD = 29.5 + 17.6 = 47.1

Pretest score boundary minimum: 12

Pretest score boundary maximum: 47

Since the selected questionnaire only allows whole number scores, the provider rounded their minimum score up (12 would be within the standard deviation range but 11 would not) and rounded their maximum score down (47 would be within the standard deviation range, but 48 would not.) For baseline and DY4 and DY5 reporting purposes, the provider will report the average most recent score of individuals whose pretest score is between 12 and 47. Individuals whose pretest scores are 11 or lower are excluded from Category 3 reporting, and individuals whose pretest scores are 48 or higher are excluded from Category 3 reporting.

Example D: Finding DY4 & DY5 Improvement Targets for Scenario 1 with a pretest score boundary

Case #	DY3	
	Pretest	Most Recent (Posttest)
1	32	35
2	23	23
3	47	48
4	40 - OOR	
5	44	46
6	41 - OOR	
7	52 - OOR	
8	7 - OOR	
9	12	12
10	31	34
11	34	37
12	17	17
13	12	12
14	33	33
15	49 - OOR	
16	24	28
17	43	43
18	15	20
19	24	24
20	31	31
21	14	19
22	58 - OOR	
23	47	47
24	9 - OOR	
25	13	18
26	39	42
27	14	14
28	39	42
29	40 - OOR	
30	41 - OOR	
31	18	18
32	50 - OOR	
33	31	31
34	48	50
35	35	35
MEAN	28.08	30.4

*OOR = Out of Range

Reporting DY3 Baseline:

In the previous example, the provider looked at their pretest scores to determine their pretest score boundary. To report their DY3 baseline, the provider then found the average most recent (posttest) score of all individuals with pretest score that was within their pretest score boundary,

DY3 Average Most Recent Score: **30.4**

To calculate their DY4, and DY5 performance targets, the provider will use the difference between the average most recent score and the average pretest score reported in DY3.

DY4 Improvement Target:

DY3 average pretest score + 1.05(DY3 Average Most Recent Score - DY3 Average Pretest Score)

$$28.08 + 1.05(30.4-28.08) = \mathbf{30.52}$$

To receive full payment in DY4, the average most recent scores of individuals who have received at least two tests, whose most recent test was in DY4 and whose pretest score was within the pretest score boundary will meet or exceed 30.52.

DY5 Improvement Target:

DY3 average pretest score + 1.10(DY3 Average Most Recent Score - DY3 Average Pretest Score)

$$28.08 + 1.10(30.4-28.08) = \mathbf{30.63}$$

To receive full payment in DY5, the average most recent scores of individuals who have received at least two tests, whose most recent test was in DY5 and whose pretest score was within the pretest score boundary will meet or exceed 30.63.

Example E: Reporting DY4 & DY5 Improvement for Scenario 1

	Case #	DY3		DY4		DY5	
		Pretest	Most Recent	Pretest	Most Recent	Pretest	Most Recent
Excerpt: cases with DY3 pretest	1	32	38		35		35
	2	23	31		23		31
	3	47	51		50		WD
	4	OOOR					
	5	44	46		48		48
	6	OOOR					
	7	OOOR					
	8	OOOR					
	9	12	12		17		26
	10	31	34		34		34
	11	34	37		37		44
	12	17	17		17		WD
	13	12	12		WD		WD
	14	33	33		36		41

Excerpt: cases with DY4 pretest	36			OOOR			
	37			24	24		35
	38			15	24		24
	39			OOOR			
	40			15	24		29
	41			23	31		34
	42			23	31		31
	43			OOOR			
	44			15	24		24
	45			32	38		WD
	46			48	48		50
	47			20	20		24

Excerpt: cases with DY5 pretest	100					OOOR	
	101					32	32
	102					OOOR	
	103					24	32
	104					34	34
	105					37	40
	106					OOOR	
	107					26	33
	108					22	22
REPORTED MEAN	28.08	32.79		33.26		33.74	
IMPROVEMENT TARGET			28.08 + 1.05(32.79 - 28.08)	33.03	28.08 + 1.10(32.79 - 28.08)	33.26	

WD = Withdrawn (the individual is no longer enrolled in the program and had no scores to report in the measurement period)

OOOR = Out of Range

In Example E, the provider exceeded their DY4 and DY5 targets. Great job provider!