maryland
health services
cost review commission

Performance Measurement Workgroup

February 17, 2021

HSCRC Quality Team

Meeting Agenda

1. COVID-19 Public Health Emergency Updates

1. Readmission Reduction Incentive Program (RRIP) Program RY 2023 PAI Follow Up

1. Other topics and public comment

RY 2022 Quality Programs: How to handle COVID for CY 2020 performance period

RY 2022 COVID Updates

- In accordance with CMS, data from January to June 2020 will not be used to assess performance
  - CMMI has indicated that they expect partial CY 2020 data to be used **unless** statistical analyses indicate that the data is unreliable
    - Jul-Dec 2020 data may be used alone, or in combination with 2019 data
  - Should Jul-Dec 2020 data be proven unreliable, CMMI **may** accept the use of RY 2021/CY 2019 performance as a proxy for CY 2020 quality performance.
- Of 41 hospitals that responded to an HSCRC survey on the use of RY 2021/CY 2019 performance as a proxy for CY 2020 quality performance:
  - All but 2 hospitals indicated that they agree for the QBR and MHAC programs
  - All but 3 hospitals indicated that they agree for the RRIP program
  - Thank you to hospitals for your Dec 2020 PMWG feedback and survey responses!
- Commission approved the removal of COVID positive patients for CY 2020; however these patients will be included in CY 2021 pending retrospective analysis
- Staff appreciate the input of PMWG and other stakeholders on potential solutions and emphasize that **there will be no perfect solution to this issue**
  - Burden of proof on 2020 data reliability is on state if no additional guidance is provided nationally

Quality-COVID Related Analyses: RRIP

- Assessment of reliability and validity of quality performance using full CY 2019 data vs. July-December 2019
  - Purpose: To understand whether 6 months of data could be ever be used; does not mean that July-December 2020 could be used without additional analyses
- Correlations of quality performance over time
  - Purpose: To understand if CY 2019 data could be re-used and considered relatively good proxy for CY 2020 performance if COVID never occurred
- Evaluation of regression adjustment for COVID cases
  - Purpose: To potentially correct performance for the effect of COVID case volume
- Report on YTD 2020 performance and discussion of combined 2019/2020 performance
  - Purpose: To assess face validity of 2020 actual performance and revenue adjustments with COVID and assess impact of combining with 2019 data to create longer time period

Readmissions Results CY 2019

Analysis	Results for Case-Mix Adjusted Readmission Rate
YoY correlations (CY13-CY19)	High correlation. Pearson correlation coefficients ranged from 0.83 to 0.93.
6 vs. 12 month correlation	High correlation. Pearson correlation of 0.96.
6 vs. 12 month rate distribution	The mean, standard deviation, and distribution of case-mix adjusted readmission rates were very similar when using 6 months versus 12 months of data.
6 vs. 12 month hospital rate	The ANOVA test comparing the mean rate when using 6 vs. 12 months of data yielded an F-value of 1.51 (p=0.23), confirming the difference is not statistically significant.
Reliability (signal to noise ratio)	High reliability overall; decreases from 0.97 to 0.91 (-5.7%) when using 6 vs. 12 months data but still high in both analyses.

Conclusions?

1. Pre-COVID readmissions over time were correlated such that 2019 data could be proxy for 2020 without COVID

1. Pre-COVID 6 month performance was correlated with 12 month performance and 6 month data was statistically reliable suggesting 6 months data could be used to assess readmissions performance; this does not mean 6 months of 2020 data during COVID is necessarily reliable due to COVID influence.

By Hospital Readmission Results 6 vs. 12 Months Data (CY2019)

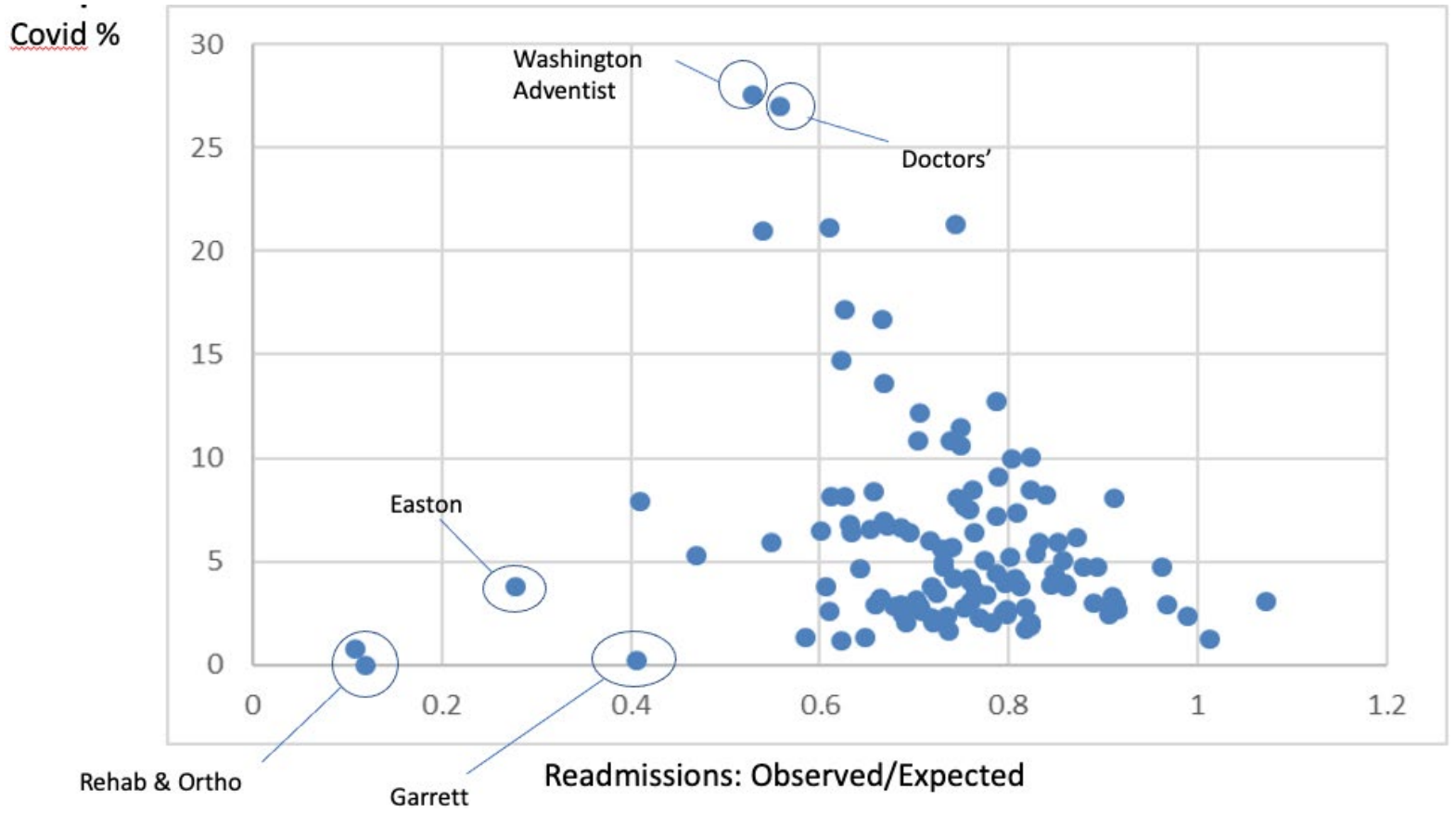
While RRIP reliability and case-mix adjusted readmissions rates change little overall, hospitals with fewer eligible discharges see more notable changes. However, this variability in small hospitals was also observed from 2018 to 2019, suggesting lower utilization has more influence on small hospital variation than seasonality.

Change in hospital-level risk-adjusted readmission rates and reliability when using 6 months instead of 12 months of data*

Metric	Max change	Avg. change	Avg. absolute change	Max % change	Avg. % change	Avg. absolute % change
Reliability	-0.26	-0.08	0.08	-38.0%	-9.7%	9.7%
Risk-adjusted rate	-1.6%	-0.1%	0.4%	-18.0%	-0.7%	3.8%

**The 6-month results were calculated using data from July 1 through December 31.*

Adjusting Readmission Performance by COVID Volume: Regression Model



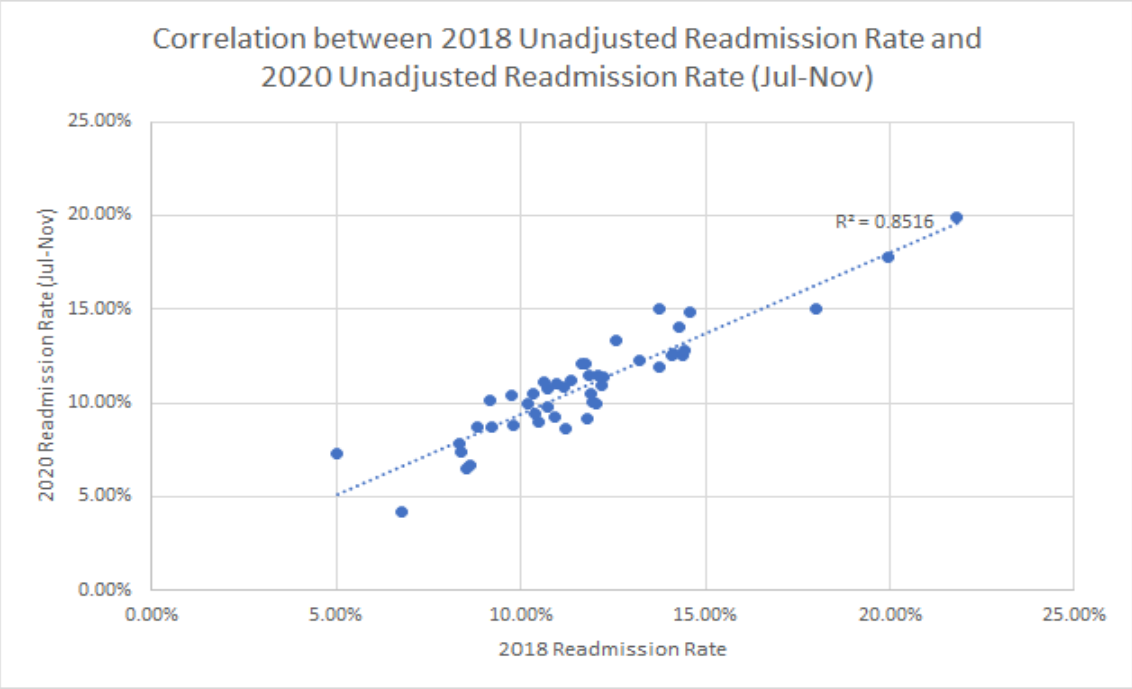
- Inclusion of COVID volume does not improve model fit
- Inverse relationship between COVID volume and readmissions gives rise to face validity concerns w/ this analytic approach and potentially evaluating readmissions generally in CY 2020

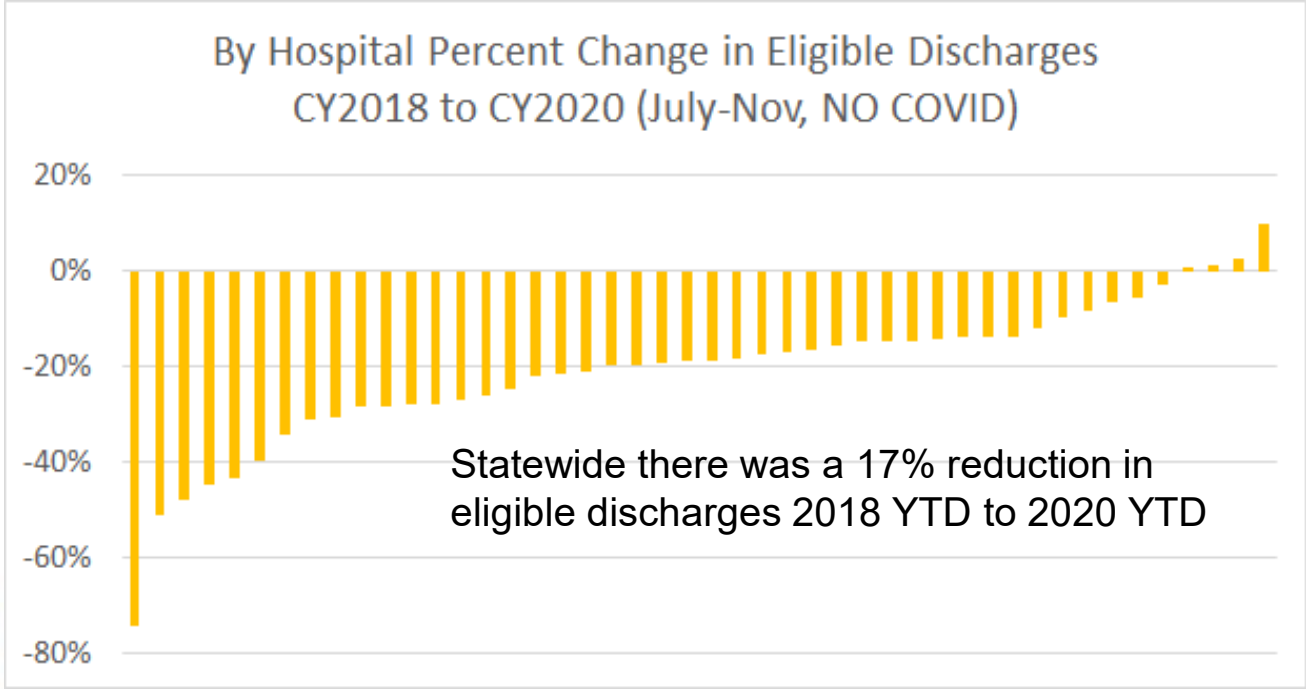
Quarterly Covid and risk-adjusted readmissions rates: April – December 2020

Readmissions 2020 YTD

COVID positive patients and Jan-June data removed

- Statewide: 9.33 percent reduction in case-mix adjusted readmission rate (5.4% simple average) versus expected reduction of 3.07%
  - Very large improvement (approximately 2/3 of what occurred in the first five years of the RRIP program) was seen for nearly all hospitals, as evidenced by material across the board reductions and very strong correlation between 2018 and 2020 YTD (0.92)
- HSCRC concerned that significant readmission reductions are a result of across the board reductions in utilization and not quality improvement



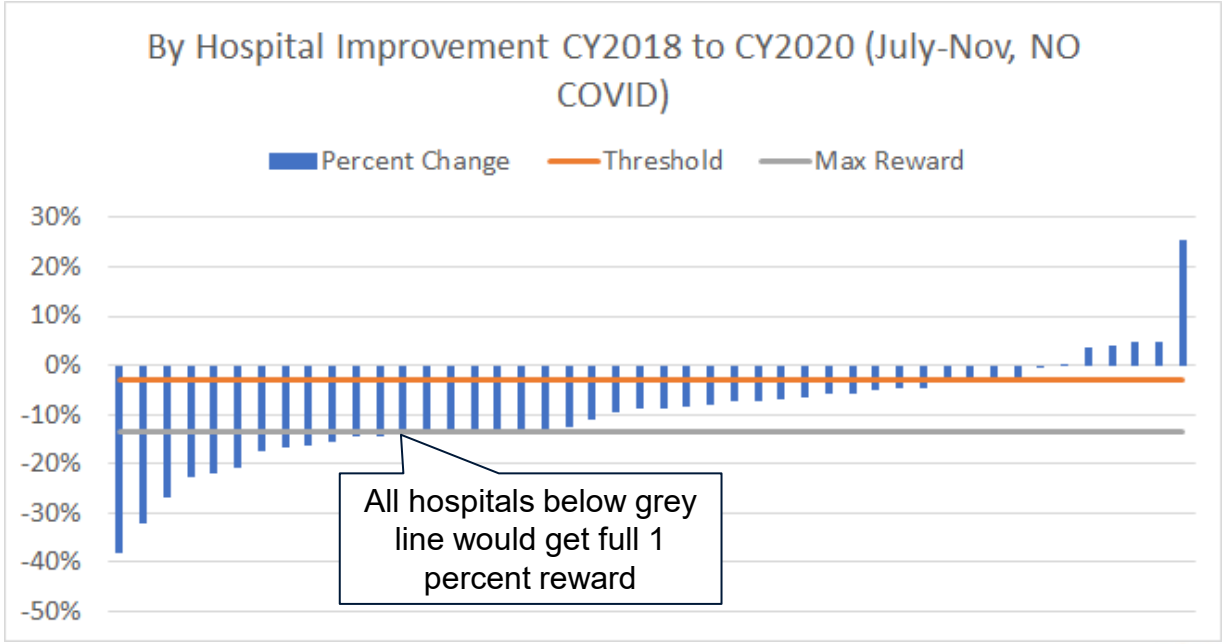


Readmission Results 2020 YTD

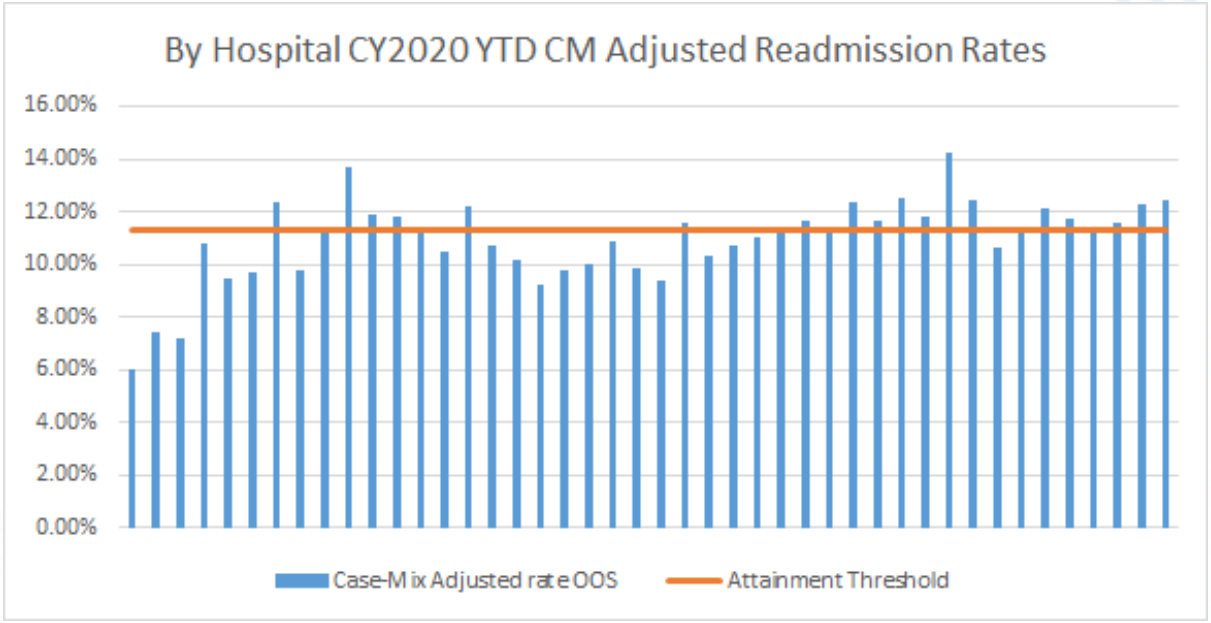
- Preliminary modeling of CY 2020 performance indicates revenue adjustments would approximately double rewards and halve penalties compared to RY 2021 assessment

- Combining CY 2019 and 2020 Jul-Nov results in lower rewards and penalties compared to RY2021 assessment; however, application of a 2020 improvement standard when assessing 2019 introduces potential concerns
  - In this modeling the improvement goal for CY 2019 + 2020 remains the same at the 2018-2020 goal (-3.07%), but 2019 represents 70% of the performance assessment

- Unclear if 2020 data is not definitively indicative of quality of care, but extent of rewards, especially relative to RY 2021, does raise validity concerns



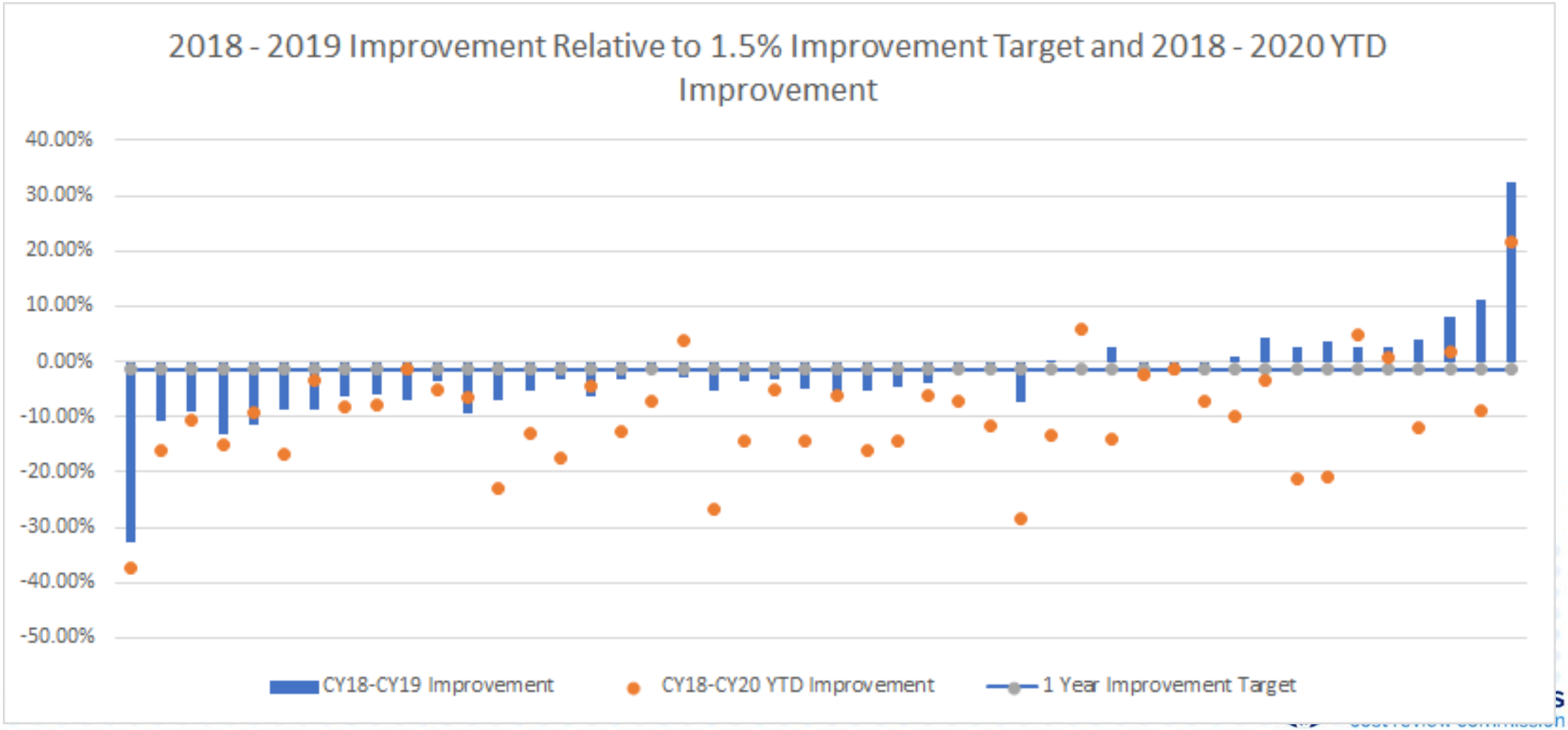
Hospitals Sorted in Same Order on Graphs



RY2021 OOS ratios used for this modeling and will be updated for final

Readmission Analysis: CY 2018 to CY 2019

- While most hospitals achieved the CY18-CY19 improvement target of -1.5 percent in 2019 (3% statewide), this did not drive the rather unrealistic improvement in CY 2020 (9% statewide)



Readmissions Discussion and Next Steps

- Thoughts on current analyses?

- Additional analyses and/or alternative methodologies?
  - Assess reliability of 2020 data (signal to noise ratio) once preliminary data through December is available
  - Follow-up with CMMI on face validity concerns

Quality-COVID Related Analyses: MHAC

- MHAC analyses more complicated due to 14 payment program measures
  - Analyses done at both PPC measure and MHAC score level

- Assessment of reliability and validity of quality performance using full CY 2019 data vs. July-December 2019
  - Purpose: To understand whether 6 months of data could be ever be used; does not mean that July-December 2020 could be used without additional analyses

- Correlations of quality performance overtime
  - Purpose: To understand if CY 2019 data could be re-used and considered relatively good proxy for CY 2020 performance without COVID

- Evaluation of regression adjustment for COVID cases Awaiting results from MPR
  - Purpose: To potentially correct performance for the effect of COVID case volume

- Report on YTD 2020 performance and discussion of combined 2019/2020 performance

MHAC/PPC Results (see Appendix for additional info)

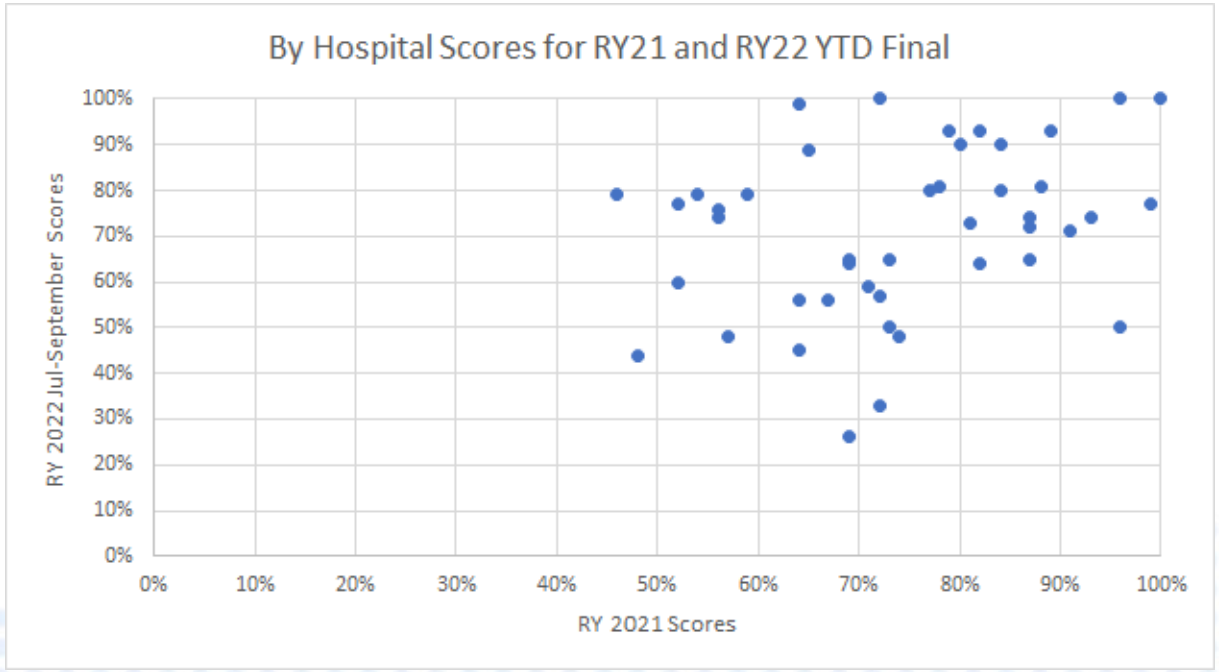
Analysis	Results for MHAC Scores	Results for Payment Program PPC Measures
YoY correlations (CY16-CY19)	Moderate correlation. Pearson correlation coefficients on the MHAC score ranged from 0.43 (RY18-19) to 0.68 (RY20-21)	Individual PPC measures showed mixed results with some PPCs having low correlation overtime.
6 vs. 12 month correlation	The correlation coefficient for the weighted MHAC score was 0.794, suggesting strong correlation between results using 6 months versus 12 months of data.	O/E ratios had moderate to high correlation (Pearson correlation coefficients ranged from 0.484 for PPC 37 (Post-Operative Infection & Deep Wound Disruption Without Procedure) to 0.946 for PPC 61 (Other Complications of Obstetrical Surgical & Perineal Wounds).
6 vs. 12 month score distribution	The mean, standard deviation, and distribution of the MHAC score were similar using 6 months versus 12 months of data.	The variation of PPC O/E ratios across hospitals increased considerably when using 6 months of data, particularly for measures with fewer hospital observations.
6 vs. 12 month hospital O/E ratio	The ANOVA test on MHAC score rejected the hypothesis that the measures/score were statistically significantly different in means using 6 vs. 12 months of data.	The ANOVA test on each the 14 PPC O/E ratios rejected the hypothesis that the measures were statistically significantly different using 6 vs. 12 months of data.
Reliability (signal to noise ratio)	Composite and Combined PPC measure* showed moderate reliability when using 12 months of data but lower reliability with 6 months data	Individual PPC O/E ratios show lower reliability when using 6 months of data even after removing small hospitals.

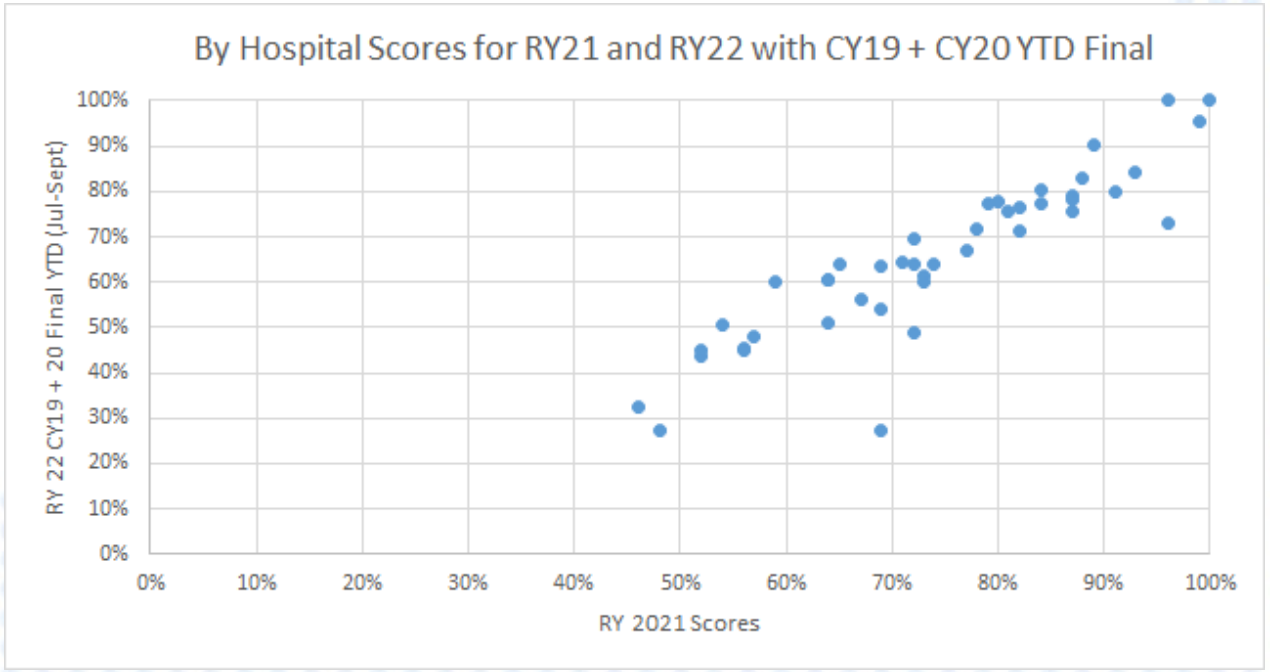
*Composite measure added up O/E ratios, combined measure created new PPC indicator of yes/no. Both measures at hospital level were correlated with MHAC score, suggesting reliability analysis is indicative of MHAC policy assessment.

MHAC 2020 Final Jul-Sept Results

Attainment only but based on pre-COVID performance standards

- RY 2022 YTD results with one quarter data result in higher penalties ($20M) and higher rewards ($51M), with net revenue adjustments lower than RY 2021 (+$31M vs +$38M)
  - Lack of correlation may be due to version differences, 1 quarter data, and COVID concerns
- Stronger correlation of RY21 and RY22 with CY19 + CY20 YTD





MHAC/PPC Discussion and Next Steps

- Thoughts on current analyses?
  - MHAC scores and PPC O/E ratios using only 6 months of data appear less reliable
- Additional analyses and/or alternative methodologies?
  - Could assess reliability of 2020 data (signal to noise ratio) using final data through December but not available until mid-April at the earliest

COVID and QBR Analyses

- Majority of the QBR score (90%) is determined using data for Hospital Compare
  - Working with CMMI to understand how HSCRC will get data for QBR with the Jan-June data excluded
  - Are there other options?

- For RY 2022 Update Factor in June, staff propose that RY 2021 be used as a placeholder
  - QBR data is always delayed such that final QBR adjustments are not implemented until Jan of RY
  - CMMI is fine with this interim approach but as with MHAC and RRIP they are wanting us to use CY2020 data

COVID and PAU Analyses

- Decisions on readmissions for PAU
  - RRIP analyses will inform decisions on the PAU readmissions measure

- Currently processing data to get 6-month and 12-month per capita PQI rates for CY 2019 and CY 2020 (preliminary data through December)
  - Will review and conduct similar correlations and reliability analyses as for RRIP
  - As with RRIP and MHAC, COVID positive patients are being removed as flagging a PQI

RY 2022 Quality Programs: Next Steps

- Concerns over using CY 2020 performance for RY 2022 are numerous:

	RRIP	MHAC
Data Reliability	Strong data reliability for 6 and 12 months but readmissions from CY 2018 to CY 2020 YTD improving by approximately ⅔ of what was achieved in 5 years of the RRIP program under the All-Payer Model strains credulity.	Data is demonstrably less reliable using 6 months of data and CY 2020 YTD performance has limited relationship to CY 2019 despite program maturity.
Face Validity of Scores	RY 2022 YTD Net Revenue Adjustment is materially greater than RY 2021 revenue adjustment, which maintained less aggressive performance standards.	RY2022 MHAC scores uncorrelated with previous performance; concerns on case-mix adjustment using historical data.
Construct Validity	Significant readmissions improvement and inverse relationship between COVID volume and readmissions suggests CY 2020 performance is not indicative of quality of care.	Utilizing CY 2019 data, as a necessity to improve reliability, that results in all but 4 hospitals with diminished performance, due to lack of relationship between CY 2019 and CY 2020, is not indicative of actual quality of care in CY 2020.

- To date the most reasonable approach to assessing RY 2022 performance is using RY 2021 revenue adjustments, but staff will continue to work through assessments to rule out any potential use CY 2020 performance.
- For the time being, staff advise the industry to use RY 2021 revenue adjustments for internal budgeting.
- PMWG will meet for COVID specific meeting in March finalize decisions on RY 2022

Readmission Reduction Incentive Program (RRIP): Patient Adversity Index Follow-Up

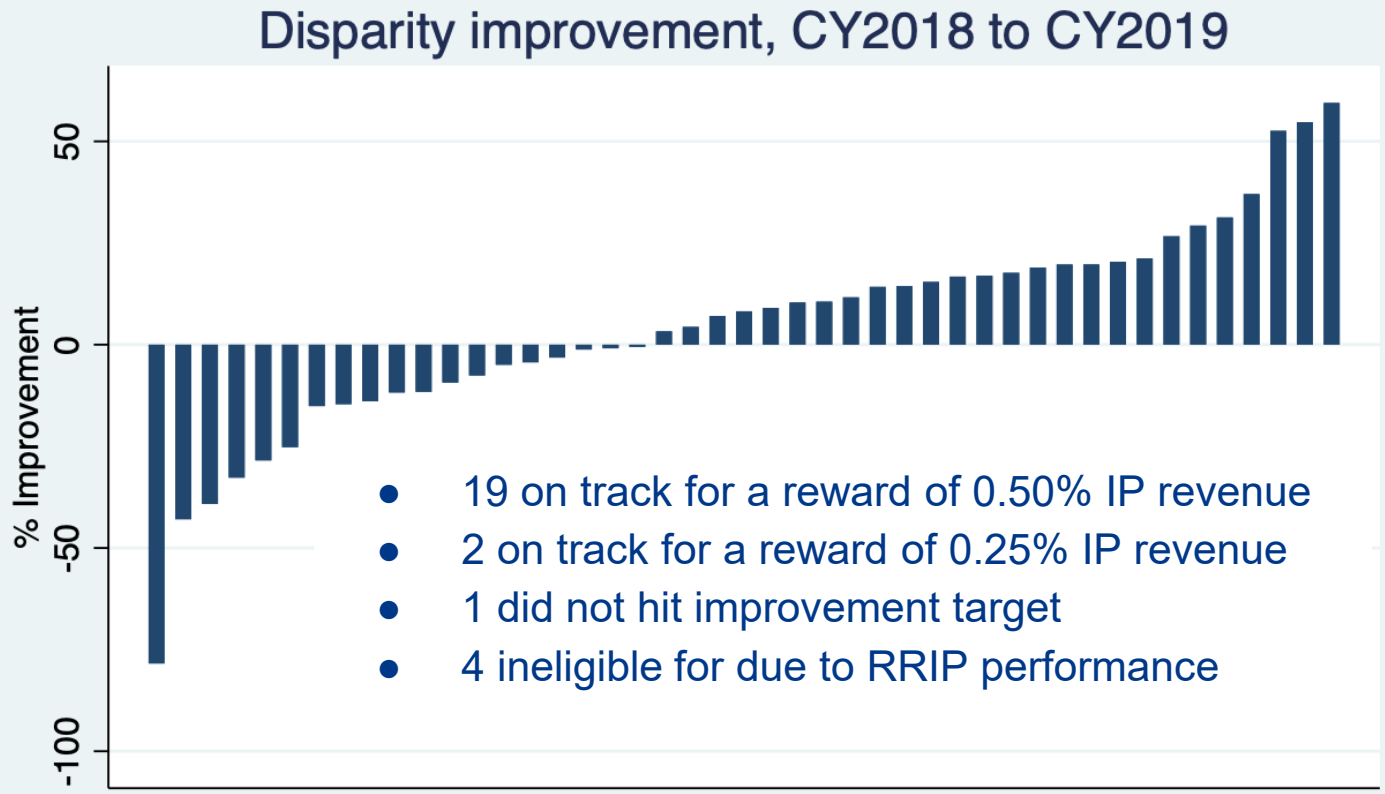
Final RY 2023 RRIP Recommendations

1. Maintain the 30-day, all-cause readmission **measure**.
    a. Remove Pediatric Oncology cases, in accordance with the intention of the oncology readmission measure.
2. **Improvement Target** - Maintain the RY 2022 statewide 5-year improvement target of -7.5 percent from 2018
3. **Attainment Target -** Maintain the attainment target whereby hospitals at or better than the 65th percentile statewide performance receive scaled rewards for low readmission rates.
4. For improvement and attainment, increase the **maximum reward** hospitals can receive to 2 percent of inpatient revenue and maintain the **maximum penalty** at 2 percent of inpatient revenue.

Final RY 2023 RRIP Recommendations (Continued)

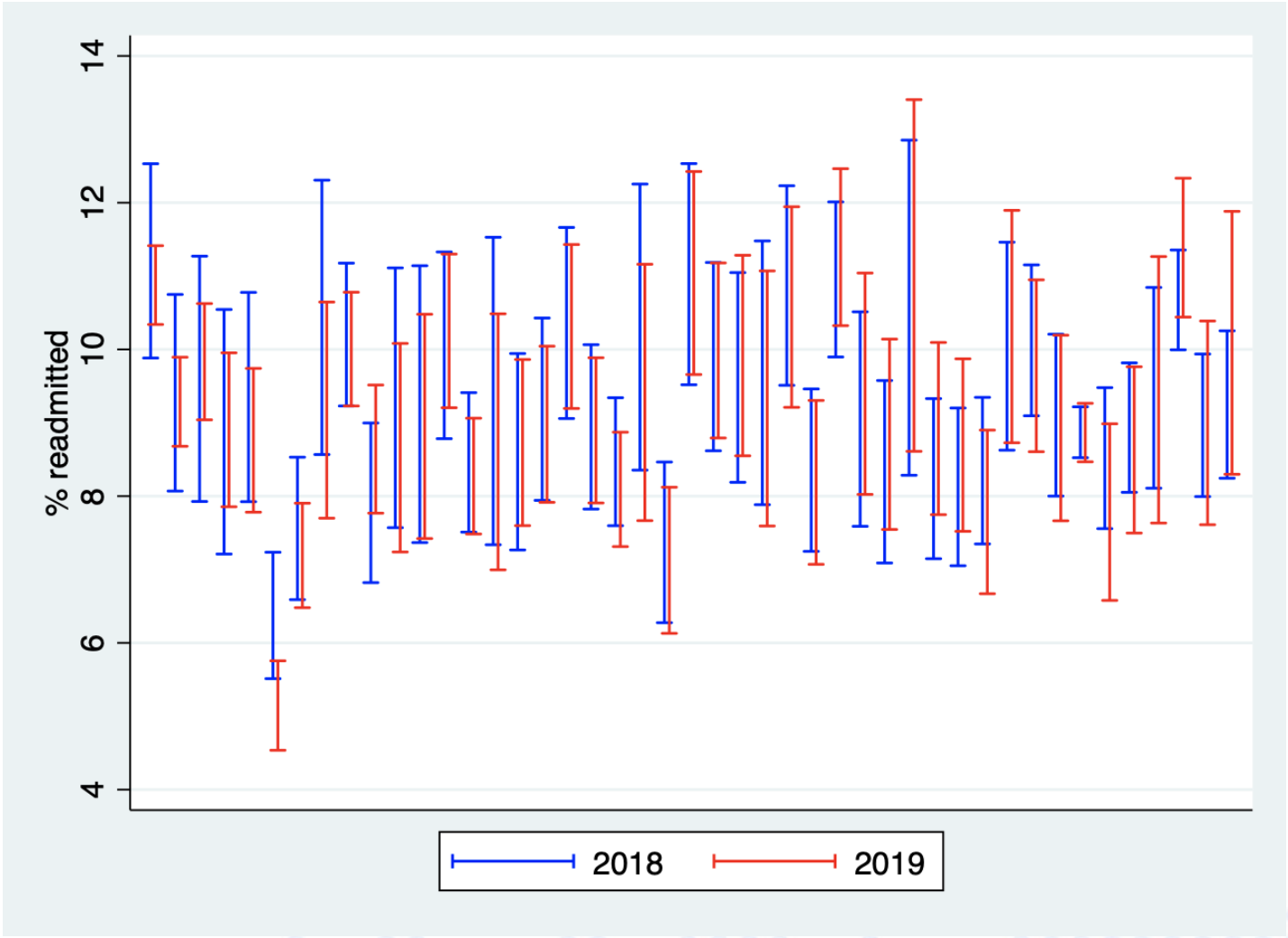
5. Provide additional payment incentive (up to 0.50 percent of inpatient revenue) for **reductions in within-hospital readmission disparities.**
    a. **Scale rewards** beginning at 0.25 percent of IP revenue for hospitals with 50 percent reduction in disparity gap measure, capped at 0.50 percent of IP revenue for hospitals with 75 percent or larger reduction in disparity gap measure.
6. Continue development of an **all-payer Excess Days in Acute Care measure**
7. Adjust RRIP due to **COVID-19 Public Health Emergency** as follows:
    a. **For RY 2022** (CY 2020 performance period) - Exclude COVID-19 positive cases; Exclude the data for January to June 2020; Evaluate what data to include; Evaluate case-mix adjustment and performance standards concerns
    b. **For RY 2023** (CY 2021 performance period) include COVID-19 positive cases but retrospectively assess any case-mix concerns, including the use of a pre-COVID time period to determine normative values.

Disparity Performance, CY 2019



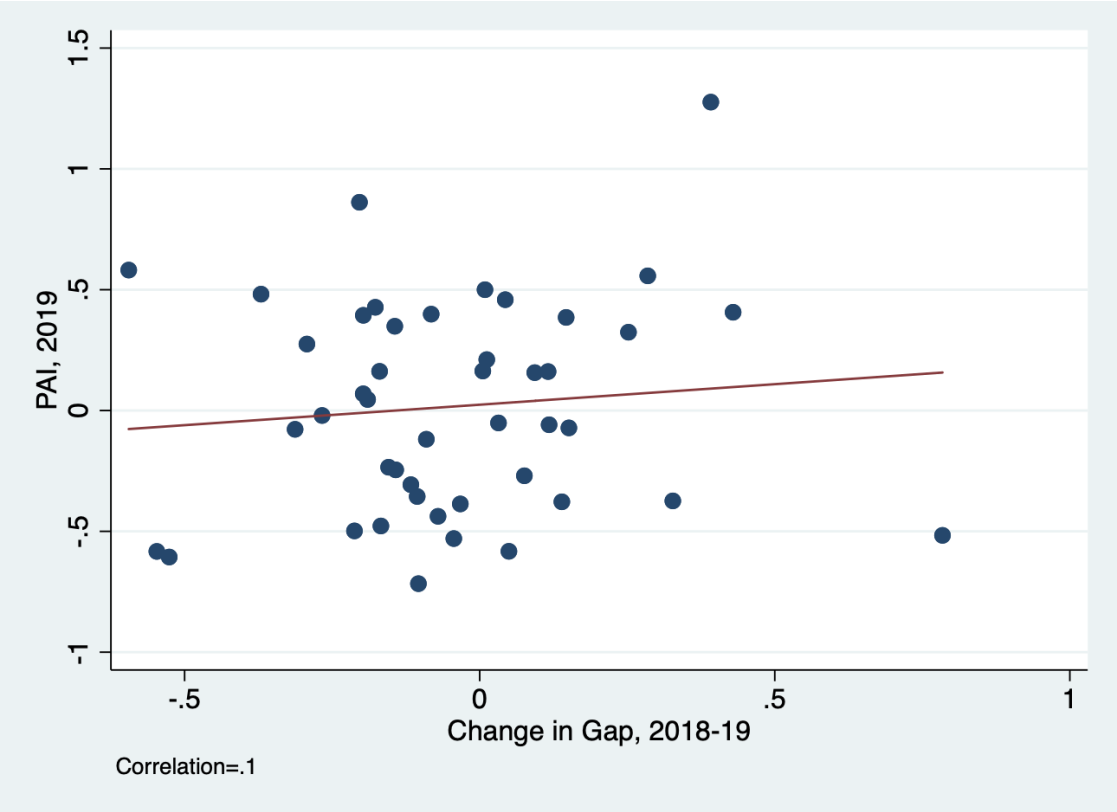
- Improvement rates higher than anticipated from RY 2022 policy due to use of actual performance period for centering of variables
- Estimated rewards of 2018-2019 improvement would be $20.3 million if policy had been in place (~half of RRIP Policy rewards)
- Thus RY2023 policy proposes increased improvement targets for rewards

Improvement Over Time



- Some hospitals saw ~50% disparity improvement. This is **not** equivalent to improvement of 50% on readmission rate
- While % change in the gap is large, changes in readmission rates for high- and low-PAI patients are smaller

Hospital Mean PAI and 2018/19 Disparity Performance

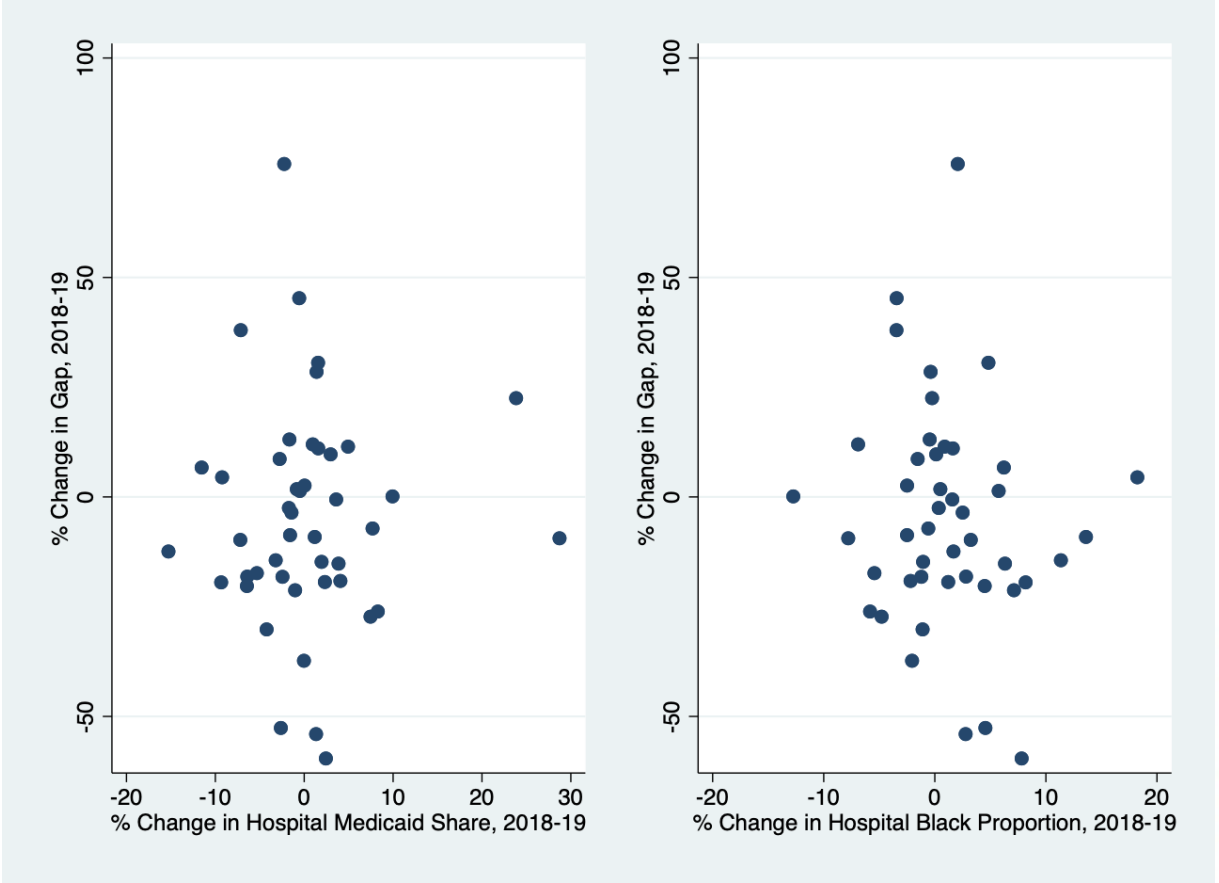


- Hospitals at all levels of PAI saw improvement on disparity measure
- 2019 PAI is not associated with improvement

Disparity Measurement

Are changes in gap driven by hospital-level shifts in race, Medicaid, ADI?

- No relationship between gap improvement and changes in race, Medicaid
- Results for ADI (not shown) are similar



HSCRC Evaluation of Race Data Quality

- Starting in 2013, in conjunction with MHA and other collaborators, HSCRC implemented training for hospitals on best practices for gathering race data. This training and related information remains available on the HSCRC website.

- HSCRC evaluated case mix race data by comparing across secondary data sets - Census and CCLF. These analyses show strong agreement at the more summarized levels (e.g. % Black was highly correlated).
  - It is difficult to assess accuracy at the hospital level for other categories, such as Asian and Hispanic/Latinx, due to small numbers. Smaller categories and “other” tend to vary due to definitional differences between sets.

Key Takeaways on Race Data Quality

Staff believe race data captured in case mix is substantially correct, both at a state and regional level, particularly in the basic categories of Black and White.

- Efforts to evaluate other data sets and continue to improve the data are ongoing.
- While the Commission should remain aware of potential accuracy issues, the Commission should not hesitate to use the data judiciously in policies.
- Staff are pursuing a number of projects that will include race data, and using the data is one of the best ways to ensure continued improvement in its accuracy.

Quality Based Reimbursement (QBR) Program

QBR RY 2023 Final Recommendations

Update on hospice for mortality

HSCRC discharge disposition codes:

50 = TO HOSPICE AT HOME

51 = TO HOSPICE DEFINED AS A MEDICAL FACILITY (CERTIFIED) PROVIDING HOSPICE LEVEL OF CARE (INCLUDES HOSPICE UNIT OF AN ACUTE CARE HOSPITAL)

Usage Note: *Include if the patient is discharged to a qualified inpatient facility and the patient will receive inpatient hospice or respite level of care. Do not use this code if the patient has not yet been accepted by a hospice and the level of hospice care is unknown.*

Other Thoughts or Questions?

Next PMWG Meeting: March 17, 9:00 AM-10:00 AM

Appendix

COVID Analyses

YoY RRIP and MHAC Correlations

Correlations calculated using RY 2022 methodology

Rate Years	Program	Metric	Pearson Correlation
2021 to 2020	MHAC	MHAC Score	0.68
2020 to 2019	MHAC	MHAC Score	0.51
2019 to 2018	MHAC	MHAC Score	0.43
2021 to 2020	RRIP	Case-mix adjusted rate	0.87
2020 to 2019	RRIP	Case-mix adjusted rate	0.83
2019 to 2018	RRIP	Case-mix adjusted rate	0.85
2018 to 2017	RRIP	Case-mix adjusted rate	0.83
2017 to 2016	RRIP	Case-mix adjusted rate	0.84
2016 to 2015	RRIP	Case-mix adjusted rate	0.93
2015 to 2014	RRIP	Case-mix adjusted rate	0.90

MHAC: Moderate correlation in scores YoY

RRIP: High correlation in readmission rate YoY

Readmissions 6-12 Month Results (CY19)

Statewide readmission rates with 6 months (10.31%) vs 12 months (10.40%) are fairly similar and both show high reliability
(see slide in presentation on max by hospital change)

Risk-adjusted readmission rate descriptive statistics

Data Period*	N	Mean	Median	Std Dev	10th Pctl	25th Pctl	75th Pctl	90th Pctl
12 months	45	0.1039689	0.1037000	0.0162393	0.0895000	0.0967000	0.1133000	0.1252000
6 months	45	0.1030711	0.1021000	0.0165856	0.0830000	0.0943000	0.1137000	0.1225000

**The 6-month results were calculated using data from July 1 through December 31.*

Reliability of risk-adjusted readmission rate

Reliability decreases 6.2% when using 6-month data

Data Period*	Overall Reliability	Median Reliability	Minimum Reliability	Hospitals with Reliability <0.5	Hospitals with Reliability <0.7
12 months	0.97	0.96	0.68	0	2
6 months	0.91	0.90	0.42	2	6

**The 6-month results were calculated using data from July 1 through December 31.*

MHAC/PPC Results 6-12 Month Results (CY19)

O/E ratios change in both directions when using 6 vs. 12 months data and some changes are quite large but overall MHAC scores does not change by much and ANOVA results say that the 6 vs 12 months PPC ratios and score are not statistically different

Descriptive statistics on hospital-level O/E ratios of payment program PPCs

		Mean			Std Dev		
PPC Number	N	12 months	6 months	% Difference	12 months	6 months*	% Difference
3	44	0.775	0.714	-7.9%	0.585	0.752	28.6%
4	42	0.704	0.643	-8.7%	0.381	0.555	45.8%
7	39	0.698	0.645	-7.5%	0.402	0.564	40.3%
9	44	0.737	0.805	9.2%	0.429	0.675	57.1%
16	38	0.656	0.497	-24.2%	0.577	0.576	-0.2%
28	27	0.926	0.765	-17.4%	0.761	0.957	25.7%
35	45	0.742	0.699	-5.7%	0.568	0.581	2.2%
37	32	1.110	0.901	-18.8%	0.622	0.803	29.1%
41	27	0.609	0.627	2.8%	0.572	0.749	30.9%
42	37	0.601	0.542	-9.8%	0.633	0.857	35.3%
49	31	0.623	0.447	-28.1%	0.565	0.686	21.4%
60	11	0.536	0.600	12.1%	0.697	0.792	13.7%
61	20	0.709	1.007	42.0%	0.756	1.127	49.2%
67	45	0.659	0.653	-0.9%	0.474	0.489	3.2%

**The 6-month results were calculated using data from July 1 through December 31.*

ANOVA test results for payment program PPCs and weighted MHAC score*

PPC Number	N	F-value	P-value
3	44	.181	.672
4	42	.344	.559
7	39	.222	.639
9	44	.315	.576
16	38	1.441	.234
28	27	.467	.497
35	45	.124	.726
37	32	1.352	.249
41	27	.009	.925
42	37	.114	.737
49	31	1.202	.277
60	11	.041	.841
61	20	.962	.333
67	45	.004	.951
Weighted MHAC score	45	.034	.853

**The 6-month results were calculated using data from July 1 through December 31.*

Descriptive statistics on hospital-level weighted MHAC score

Performance period*	N	Mean	Std Dev	Median	25th Pctl	75th Pctl
12 months	45	0.735	0.146	0.730	0.640	0.840
6 months	45	0.741	0.150	0.760	0.670	0.830
% Difference		0.8%	2.9%	4.1%	4.7%	-1.2%

**The 6-month results were calculated using data from July 1 through December 31.*

MHAC Scores 6-12 Month Results (CY19)

Reliability substantially decreases when using 6 months data

Estimated composite PPC mean reliability using 12 months versus 6 months of data

Measure	N	6 months	12 months	Change in Reliability
Composite PPC measure	45	0.523	0.690	-24.2%
Combined PPC measure	45	0.490	0.658	-25.5%

**The 6-month results were calculated using data from July 1 through December 31.*