

Rationale for adopting HRG4+ phase 3

Published by NHS England and NHS Improvement

July 2016

Contents

- Background
- Options for currency for 2017 to 2019
- Criteria for choosing currency
- HRG4+ technical performance
- Other benefits of HRG4+
- Summary

Final decision for 2016/17 was to use HRG4, prices and top ups from the Enhanced Tariff Offer

Decided not to implement HRG4+ or new Prescribed Specialist Services based top-ups.

Rationale

- Package including HRG4+, new prices and other policies was producing unexpected results, particularly not reflecting complexity or speciality as anticipated. These proposals were therefore not adopted.

New currency design

- The majority of tariff is set through Health Resource Groups (HRGs) for inpatient care (~£19bn).
- HRGs are developed by the HSCIC over a number of years, first introduced into reference costs, then can be adopted into tariff.
- HRG4+ is the latest version, with phase 3 used for 2014/15 reference costs.

How do we choose which currency design to propose going forward?



- Technical performance – does it group patients together better into like-cost groups?
- Use the most recent cost data.
- Clinically supported.
- Minimise disruption to the sector.
- Deliverability by NHS Improvement and NHS England.

Options considered for proposed 2017 to 2019 tariff



- Retain existing HRG structure
- Use HRG4+ phase 2
- Use HRG4+ phase 3

Theoretically, HRG4+ is better than HRG4



HRG4

FZ67A	Major Small Intestine Procedures 19 years and over with CC
FZ67B	Major Small Intestine Procedures 19 years and over without CC



HRG4+

FZ67C	Major Small Intestine Procedures, 19 years and over, with CC Score 7+
FZ67D	Major Small Intestine Procedures, 19 years and over, with CC Score 4-6
FZ67E	Major Small Intestine Procedures, 19 years and over, with CC Score 2-3
FZ67F	Major Small Intestine Procedures, 19 years and over, with CC Score 0-1

HRG4+ is more discriminatory as it includes a more detailed breakdown by complexity scores better reflecting the differences in costs between HRGS

Testing the theory – Technical Performance



- We have looked at the HRG design to test:
 - Does it better explain overall variation in patient level costs (Reduction in Variation)?
 - Does it reduce the variability in costs within HRGs (Coefficient of Variation)?
 - Reduction in variability between providers for each HRG (Data Variation Index).
- According to all three measures, HRG4+ phase 3 performs best, then HRG4+ phase 2, then HRG4.

Reduction in Variance (RIV)



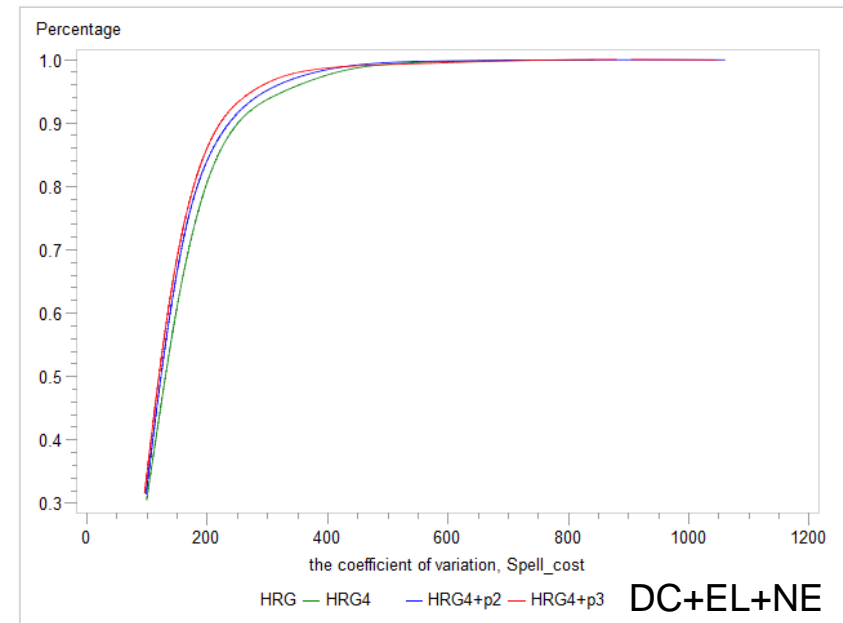
- We want the currency to explain as much of the variation in individual patient's costs as possible. The higher the RIV score, the better.

HRG	PLICS dataset	Grouper	DC RIV	EL RIV	NE RIV	DC + EL + NE RIV
HRG4	PLICS 1415	Local payment grouper 15/16	16.1%	30.0%	24.9%	25.9%
HRG4+ phase 2	PLICS 1415	Engagement grouper 16/17	16.2%	30.7%	28.1%	29.2%
HRG4+ phase 3	PLICS 1415	Reference cost grouper 14/15	17.3%	35.4%	31.0%	33.0%

- HRG4+ phase 3 explains significantly* more variation in total PLICS costs than HRG4+ phase 2.

Coefficient of Variation (CV)

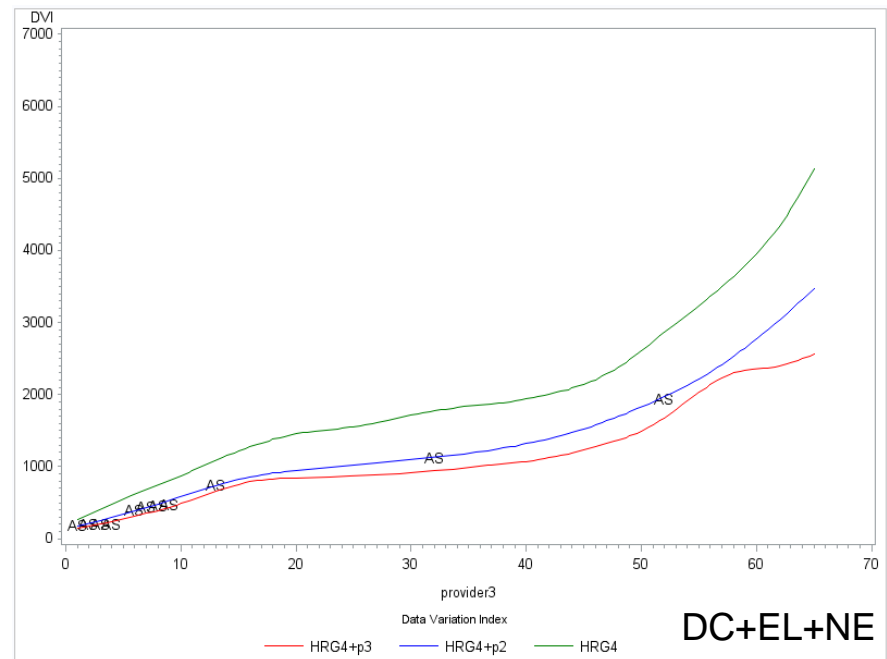
- We want to reduce the variation within each individual HRG*.
- A more convex slope implies HRG designs reach 100% coverage at lower CV scores. The higher the line, the better.
- HRG4+ phase 3 has a lower % of high variability HRGs than HRG4+ phase 2 in each point of delivery and overall.



Data Variation Index (DVI)



- DVI constructs an index of how different a provider's cost is from the average within an HRG*.
- Lower DVI means less variation within HRGs at provider level.
- Under HRG4+ phase 3, HRGs will be more consistent at provider level as the cases within each HRG are more similar in cost under HRG4+ phase 3 than under HRG4+ phase 2.



Conclusion of technical assessment

- The RIV results shows HRG4+ phase 3 explaining 3.8% more variation in total PLICS costs than HRG4+ phase 2, and 7.1% more variation than HRG4.
- The CV plots shows that across a range of highly variable HRGs there was an improvement in HRG4+ phase 3.
- The DVI showed that there was a systematic reduction in variance for HRG4+ phase 3 across HRGS at trust level when compared to HRG4+ phase 2 and HRG4.
- **All three technical measures showed that HRG4+ 3 was a better designed currency.**

Most recent reference costs are collected on HRG4+ phase 3



- Tariff isn't just about capturing clinical practice, but setting appropriate prices, capturing changes in cost.
- Option 1: Rollover – costs from 2011/12.
- Option 2: HRG4+ phase 2 – costs from 2013/14.
- Option 3: HRG4+ phase 3 – costs from 2014/15.
- HRG4+ phase 3 has the most recent and up-to-date cost data.

Further considerations still leaves HRG4+ phase 3 as our preferred option



Given current information and feedback; HRG4+ phase 3 is:

- Clinically supported – EWGs designed the HRG4+ phase 3 to better reflect clinical services at a more granular level.
- Minimises disruption to the sector – Work by Impact Assessment shows that the phase 3 prices provide a smaller disruption to the provider section than phase 2.
- Is deliverable by NHS Improvement and NHS England.

Summary

- HRG4+ phase 3 is our preferred option for the 2017 to 2019 currency because:
 1. It performs technically better (measured by RIV, CV and DVI).
 2. It is based on the most recent costs.
 3. It is supported by clinical experts.
 4. It produces less disruption to the sector than HRG4+ phase 2.
 5. It is possible to implement.

APPENDIX

1.1 RIV methodology



- RIV statistic is used to measure the explanatory power of casemix systems, i.e. the proportion of total LoS variation explained by the groups.
- A value of 0% means that the classification explains none of the variance in the dependent variable (e.g. LoS or cost), whilst 100% means it explains all of the variance. 100%, whilst theoretically possible, would suggest that all the data in each group have the same LoS/cost. Typical results for LoS would be 30-40% whilst cost would be 60-70%.
- The RIV, often expressed as R² to describe the predictive validity of the classifications, is calculated to describe the explanatory power of the grouping classifications. The unadjusted form of the calculation of RIV is the inverse of the ratio of the whole sum of squares (WSS) and the total sum of squares (TSS), expressed as a percentage.

$$R^2 = 1 - \frac{WSS}{TSS}$$

$$WSS = \sum_{j=1}^k \sum_{i=1}^{n_j} (x_{ij} - \bar{x}_j)^2$$

$$TSS = \sum_{j=1}^k \sum_{i=1}^{n_j} (x_{ij} - \bar{x})^2$$

Where

WSS = whole sum of squares

TSS = total sum of squares

k = the number of groups

n_j = the number of cases in group j

x_{ij} = value of case i in group j

\bar{x}_j = mean of group j

\bar{x} = overall mean

1.1 RIV results



- We want the currency to explain as much of the variation in individual patient's costs as possible. The higher the RIV score, the better.
- HRG4+ phase 3 explains significantly more variation in total PLICS costs than HRG4+ phase 2.

HRG	PLICS dataset	Groupier	DC RIV	EL RIV	NE RIV	DC + EL + NE RIV
HRG4	PLICS 1415	Local payment grouper 15/16	16.1%	30.0%	24.9%	25.9%
HRG4+ phase 2	PLICS 1415	Engagement grouper 16/17	16.2%	30.7%	28.1%	29.2%
HRG4+ phase 3	PLICS 1415	Reference cost grouper 14/15	17.3%	35.4%	31.0%	33.0%

- Testing the results (Chow)

HRG	DC + EL + NE F statistic	DC + EL + NE N	DC + EL + NE DF	DC + EL + NE P-value
HRG4 vs HRG4+ phase 2	0.4845501	10786741	10780897	0.05497548
HRG4 vs HRG4+ phase 3	0.7108137	11358886	11351834	0.050113252
HRG4+ phase 2 vs HRG4+ phase 3	4.3270193	10986429	10980589	0.054993894

The Chow test is a statistical and econometric test of whether the coefficients in two linear regressions on different data sets are equal.

The model in effect uses an F-test to determine whether a single regression is more efficient than two separate regressions when the data comes from two sub-samples. This gives a significance test to determine if the R^2 of two data sets are equal or not.

1.2 CV methodology



- Whilst the RIV statistic gives a result to be applied across groups, a statistic is required to measure the within-group variability or homogeneity. The ratio of standard deviation (SD) to the arithmetic mean of a group, or CV gives a measure of the relative variability within a single group.
- The CV is reported for a group to describe its homogeneity. A value of 0 would indicate that a group has no variance from the mean (i.e. standard deviation is equal to 0), whilst a CV value for a group above 1.00 would indicate heterogeneity within the group, where the standard deviation is greater than the mean. We would anticipate that the more homogenous an HRG is, the more likely it is that the underlying HRG Design is robust. Caution should be used when the mean is close to zero as CV may be sensitive to small changes in the mean.

$$CV = \frac{SD}{\bar{x}} = \frac{\sqrt{\sum_{i=1}^n (x_i - \bar{x})^2 / n}}{\bar{x}}$$

Where SD = standard deviation of the group
 = mean of group
 = value of case i in the group
 n = the number of cases in the group

1.3 DVI methodology



- The Department of Health recently described the calculation of a Data Variation Index. The DVI for a provider would be of the format below, which describes an average of the absolute percentage deviations for each individual HRG from the national average for that HRG.
- This would aim to describe the provider's difference relative to the mean. Lower DVIs at provider level might imply better defined reference costs. A systematic decrease in all DVIs across all trusts may be expected, as well as more homogenous DVIs within groups of 'similar' trusts. While this is a relatively simple calculation, it has the drawback of providing a headline figure only. This measure would also be strongly influenced by data such as extreme outliers that may lead to an inflated mean cost if they are not excluded from calculations.

$$DVI = \frac{\sum_{i=1}^n \left(\frac{|R_i - A_i|}{A_i} \times 100 \right)}{n}$$

Where R_i = the provider average cost for HRG i
 A_i = the national average cost for HRG i
 i = the list of n total HRGs