Attachment W

Distribution of CAHPS and HEDIS Scores and Assignment of Anecdotes

Distribution of CAHPS and HEDIS Scores

To reduce variability across respondents and make it easier to detect the impact of complex choice sets, we will present all respondents facing small (six doctors, four health plans) and larger (twelve doctors, eight health plans) with the same distribution of CAHPS and HEDIS scores. For each respondent, the name/gender of the doctor will be randomly assigned among the available scores. The key question is therefore how to generate the distribution of CAHPS and HEDIS scores.

Our review of the literature suggested that HEDIS-type clinical performance measures, which are generally based on dichotomous items, tend to correlate at levels very close to zero, unless one disattenuates for low reliability. This near-zero correlation is observed even for composite measures. Accordingly, clinical composites in the real world will generally appear to be uncorrelated from a consumer's perspective. Given this state of the world, we will generate simulated clinical composite scores that reflect average correlations of zero among the four clinical composites (heart care, asthma care, diabetes care, and testing/screening). For the CAHPS composites, an average correlation of 0.5 among the composites would be a good approximation of the real world.

For purposes of the experiment, it is highly desirable to present subjects who receive both CAHPS and clinical measures with a comparable amount of variation in both types of measures, both for the overall roll-up scores and for the individual composites on the drill-down. Since all scores will be presented in aggregated form using symbols, this means essentially having the same proportions of doctors receive 5 stars, 4 stars, etc. We had tentatively agreed on a 10-20-40-20-10 distribution for CAHPS scores.

Putting all this together, we suggest the following steps for generating scores:

1a. Generate joint distributions of four CAHPS composite scores that reflect underlying correlations of 0.5 among composites, with each composite distributed in the proportions described above.

1b. Generate joint distributions of clinical composite measures that reflect underlying correlations of 0 among composites, with each distributed in the same proportions as CAHPS composites.

2a. For the CAHPS roll-up measure, add up the number of stars across composites. In the resulting distribution, assign cut points to create five levels with a distribution that is as close as possible to the distribution of stars for underlying composites. More precisely, we'd assign median percentiles for each of the star categories, so that 1 star = 5 points, 2 stars = 20 points, 3 stars = 50 points, 4 stars = 80 points, and 5 stars = 95 points.

2b. Carry out the parallel step to 2a for the clinical measures. Note that this will yield a different distribution of summed points than is obtained for the CAHPS measures, because of the zero correlation among clinical composites versus the 0.5 correlation among CAHPS composites. In the resulting distribution, assign cut points to create five levels on the roll-up measure with a distribution as close as possible to the distribution for the CAHPS roll-up measure. (This will require a different set of cut points from the ones chosen for the roll-up distribution of CAHPS scores).

Note that this process should yield comparable levels of variability in CAHPS and clinical roll-ups and CAHPS and clinical composite measures. Where the CAHPS and clinical measures will differ is in the relationships between the roll-up and drill-down results at the composite level. A 5-star physician on the CAHPS roll-up measure will tend to have 4 or 5 stars on the underlying composites, whereas a 5-star physician on the clinical roll-up will tend to be more inconsistent in the number of stars on specific composites (e.g., more 3-star results). We think this is an interesting byproduct of the greater consistency of CAHPS composites compared to clinical measures, and that it reflects what consumers are likely to encounter in the real world.

Assignment of Anecdotes Based on CAHPS Scores

To be realistic, we wanted the general tenor of the anecdotes to be more positive for clinicians/plans that had higher CAHPS scores, but with considerable variance in this relationship. We therefore grouped anecdotes to produce a modal emotional valence of highly negative, mildly negative, mixed (evenly positive and negative), mildly positive and highly positive, which involved different combinations of statements based on the emotional valence that they were assigned in pilot testing.

For each respondent, each of the clinicians in their choice set will first be randomly assigned a modal emotional valence, such that these modal valences have a 0.35 correlation with their aggregated CAHPS score. (See Exhibit 1). Following this protocol, a clinician/plan with a one star (lowest consumer rating) has a 69.6% probability of being assigned anecdotes with a generally negative orientation, but some small probably of receiving more generally positive anecdotes. (This variation will allow us to explore than anecdotes contradicting CAHPS scores dominate those scores influence consumer choice).

Once a clinician/plan has been assigned a modal affect, a second random draw will determine the number of anecdotes (ranging from 4-6) that they are assigned, which will in turn determine the mix of the emotional valences for the individual anecdotes (Exhibit 2). Reflecting patterns observed on actually physician commentary web sites, even physicians with a generally negative or generally positive reputation will have at least one comment that runs contrary to this general pattern.

EXHIBIT 1

Probability Any Clinician is Assigned A Modal Affect for the Associated Anecdotes

	Strong Negative	Weak Negative	Mixed	Weak Positive	Strong Positive	
AGGREGATE CAHPS Score	1	2	3	4	5	
	……..	……..	……..	……..	……..	……..
1 One star	**32.7%**	26.9%	26.0%	10.4%	4.0%	10% fixed
2 Two stars	21.7%	24.8%	**30.4%**	15.4%	7.7%	20% fixed
3 Three stars	13.6%	20.6%	**31.7%**	20.6%	13.6%	40% fixed
4 Four stars	7.7%	15.4%	**30.4%**	24.8%	21.7%	20% fixed
5 Five stars	4.0%	10.4%	26.0%	26.9%	**32.7%**	10% fixed
	15% expected	20% expected	30% expected	20% expected	15% expected	
…….....	……..	……..	……..	……..	……..	……..
	Row %s are entries					r=0.35

Exhibit 2

Number of Comments With Each Type of Valence

Modal Valence	Strongly Negative	Mildly Negative		Mildly Positive	Strongly Positive	Number of Comments
Strong Negative	3	2		1		6
	2	2		1		5
	2	1		1		4
Mild Negative	1	2		1		4
	1	3		1		5
	2	2		2		6
Mixed		2		2		4
	NA	NA	NA	NA	NA	5
	1	2		2	1	6
Mild Positive		1		2	1	4
		1		3	1	5
		2		2	2	6
Strong Positive		1		1	2	4
		1		2	2	5
		1		2	3	6