Updates

2001

September 7

2000

September 22

August 29

June 1

April 3

1999

November 24

October 13

September 3

July 23

February 15

1998

November 18

October 6

September 25

September 18

September 9

Update # 13

September 18, 1998

Responding to the feedback received regarding both Phase II reports

The Commission received many submissions as a result of both Phase II reports which detailed the methodology to be used in developing a RBRVS. We have spent much time considering the comments and concerns that were raised. Some aspects of the methodology have been refined to incorporate the constructive input we received.

The following responds to some of the most common questions and comments we heard. In reviewing this document, it may be helpful for you to consult the second Phase II report (dated May 1998).

We will be communicating with a number of the Sections prior to the reference code session to ask them to expand on their submissions or to respond to specific questions they have raised pertinent to their Section or speciality. If you would like to meet with the Commission or discuss your submission further, please call the Commission with your request**.**

Pilot study *(see pp.18-37 of May 1998 report)*

> ***Question:***How does the Commission respond to the low response rate in recruiting for the pilot study? Could the response rate have invalidated the study itself? How could such a low response as well as the participation of only four OMA Sections guarantee the participants were representative?
>
> ***Answer:*** Our experts tell us that the response rate was not a concern in this case because of the design and purpose of the study.In addition, the participants in the study were randomly allocated to three groups (global, composite/own and composite/mixed) and the key questions revolved around differences among the groups in the way they valued the codes. Thus, the aim was not to generalize the membership of the OMA as a whole, but to see if different valuing techniques produced different outcomes in terms of the reliability of the total ratings.
>
> In any case, Decima Research, the firm that managed the recruitment for the pilot study, noted that the study and/or the Commission, were not given as reasons for non-participation.

July 3

May 29

April 8

March 30

February 4

1997

November 19

October 23

September 4

August 7

July 29

July 2

Question: Why didn't the pilot study evaluate all aspects of the methodology?

Answer: The purpose of this pilot study was to evaluate key components of the methodology. It would have been extremely difficult logistically to examine all possible combinations of each possible approach in one pilot study.

The results of the pilot study were used to guide our thinking - just as the consultations we continually have with physicians and experts, the focus groups we've conducted, and our review of the scientific literature are also used to guide our thinking. The final methodology has emerged over time and takes into consideration the input and insight we've received.

Question: Why is the consensus approach not being used?

Answer: With the global method, the consensus ratings were consistently higher than the averages of the individual judgements. Across the pilot groups, the dynamic was that groups achieved consensus by all agreeing to higher ratings, despite the lower ratings given to those same services when they were evaluated by the members of the groups each evaluating services on their own.

Individual physicians to evaluate survey codes *(see p. 48 of May 1998 report)*

> ***Question:*** Won't the values given to the services be based on subjective judgements? Isn't this supposed to be a scientific exercise?
>
> ***Answer:*** Almost every aspect of the RBRVS is based on the expert external determination, collection and analysis of objective validated data - i.e. the time study, as well as the derivation and application of practice costs and lost opportunity costs. Determining the intensity of a service requires a different approach, one which relies on the subjective judgement of many individuals. However, in this case, 'subjective' is not synonymous with 'unreliable' or 'invalid'.Our experts tell us combining subjective ratings across many evaluators can result in highly reliable and valid results. The pilot study and other more informal studies we've conducted since then have confirmed that. The Commission is confident that the methodology that is being used to evaluate the intensity of the selected reference and survey codes is both sound and reliable.
>
> The evaluation of the intensity of the survey codes will be done by hundreds of physicians across the province. We've asked each Section to provide the Commission with a Section profile to assist in the recruitment of physicians from their speciality. The Angus Reid Group will manage the recruitment of a broad representation of physicians from each speciality. This will strengthen even further the reliability and validity of the outcome.

Times for survey codes to be provided *(see p. 48 of May 1998 report)*

> ***Question:*** Will providing time data to those evaluating the intensity of each survey code not unduly influence and possibly invalidate their evaluations?

Answer: This question was raised by members of the National Advisory Committee as well as the Commission's methodology expert panel. The Commission has reconsidered providing the time of providing a service as part of the service code descriptor for the evaluation of intensity. Physicians' evaluations of time are being gathered through the Commission's time survey, separate from physicians' evaluations of intensity to be gathered in the reference code session (in the May report, referred to as the "cross sectional working group") and survey code evaluations. All of the resource inputs will be factored together after all the inputs have been evaluated and validated.

Weighting among the intensity resource inputs *(see p. 50 of May 1998 report)*

Question: What is the Commission's rationale for the inclusion of communications and interpersonal skills as an input, and for an equal weighting of all the intensity inputs?

Answer: Some Sections have stated that communications and interpersonal skills are related to the physician, not to the service, and are therefore inappropriate as a resource input. The Commission has considered this issue very carefully. Certainly the individual personality of a physician and his or her *personal* communications attributes are important. However, this is not what this input is meant to address. It is the view of the Commission that medical services require varying levels of *professional* communications and interpersonal skills from the physician providing the service. Not personality, but professional learned skill directly related to the particular service being provided.

Some have commented that because communications and interpersonal skills have not been a separate input in previous RVS exercises, it should not be in this one as well. The Commission believes this has been an omission that required correction.

The related question most often heard is why should the intensity inputs be weighted equally? While that question is raised frequently there is absolutely no agreement on an alternative approach. Which intensity input is of more value than the others? A medical service may require less of one input and more of another. Another medical service may be absolutely reliant on yet another input.

The Commission is of the view that in the absence of compelling data, evidence or agreement to do otherwise, there should be equal weighting given to the four intensity inputs. The Commission sees the range and diversity of services across all medical specialities as simply too large to set a differential value for each input and make that value uniform across all medical services without strong data or agreement among the specialities.

The Commission's view is that the evaluation process we are about to undergo, using individual physicians evaluating their own services, should determine - service by service - the value of each intensity input relative to the others. The

comparative value of each intensity input should be reflected in the rating given to each input by the evaluators of that service.

Weighting time versus intensity *(see p.51 of May 1998 report)*

Question: Why was your preliminary decision on the weighting of time versus intensity that it should be 4:1?

Answer: Quite simply, because it seems to make the most sense. The Commission consulted widely on the question of how time and intensity should relate to each other. We asked the question: How much should a unit of time at the highest intensity of work be valued against the same unit of time at the lowest intensity of work? We met with our expert panel as well as recognized experts in job evaluation methodology, discussed this question with representatives of other professions, looked at the U.S. approach and other precedents and consulted with the National Advisory Committee and with the OMA Sections. These consultations and analysis seemed to support a weighting of 4:1.

Clearly, there is no prevailing gold standard to use to determine this issue and the Commission has been given no significant data or evidence to support another weighting approach. The Commission is well aware of the importance of this issue and intends to undertake further analysis and additional consultation before finalizing the weighting of time versus intensity.

Allowance for global adjustment of total rating *(see p.51 of May 1998 report)*

Question: Does allowing each evaluator of the survey codes to review the results of their individual intensity input evaluations and to make changes to each service's total rating invalidate an otherwise relatively objective evaluation exercise by adding a subjective judgement call?

Answer: The global adjustment was conceived as a way to allow evaluators a subjective, overall consideration of their evaluations. The Commission was looking for a method that provided for a consistent and prescribed method of evaluating the intensity inputs, but one that also allowed evaluators to be fully comfortable with their ratings of intensity. We wanted evaluators to be satisfied with the each part - i.e. their evaluations for each intensity input - as well as with the total intensity, understanding that the four inputs rated separately might not adequately recognize an individual's implicit judgement of the overall intensity of a service. We recognized that this concept was difficult to incorporate into a methodology that was intended to be as objective and reproducible as practicable, but we felt that the allowance for a tightly prescribed global adjustment might accommodate our aim.

Subsequent to the Phase II reports we heard the following:

i. We were advised by the experts we consulted that allowing evaluators to make global adjustments to ratings that included time was not optimal

from a methodological point of view. Their concern is that this might implicitly change the value of time – data which had been gathered and validated in a separate process. Rather, the gathering and validation of the time data for each service should be kept completely separate from the intensity evaluations for that service.

ii. Some Sections told us that a global adjustment to the total intensity score would invalidate the rating given to each individual intensity input.

iii. No Section strongly supported the concept of the global adjustment as presented.

As a result, we have decided:

i. The composite analysis of intensity on its own is too rigid. Some additional inclusion of flexibility and allowing evaluators to reflect carefully on their ratings should be included.

ii. If global adjustment is to be allowed on total intensity, it should not include time or weighting. These will be kept separate from the intensity evaluations.

iii. With the guidance of its experts both in methodology and in evaluation techniques, the Commission will test, through focus group testing as well as by a pre-test of the evaluation method, how best to apply or revise the concept of global adjustment. (i.e. do physicians find it useful to their evaluation of services? How could it be better incorporated into the methodology?)

iv. Based on that qualitative input, either the concept of global adjustment will be altered to apply to total intensity only (i.e. not including time or weighting), with strict parameters for how and in what circumstances it can be used, or it will be not be included as part of the evaluation of total intensity. Instead, individuals will be permitted to adjust their evaluations of each intensity input (again, within limits) to meet the goals of flexibility and individual judgement.

v. If the global adjustment is not included, the Commission will ensure that, without comprising the overall methodology, the process for evaluating the intensity inputs provides evaluators with flexibility to reflect on their evaluations of each service and make appropriate changes.

vi. Whatever the details of the final methodology, reproducibility and transparency will be important guiding principles.

Integration of evaluated codes *(see p. 52 of May 1998 report)*

Question: How will the Commission determine what linking is required, and, if so, by what method?

Answer: Linkage across Sections is a double-edged sword. On the one hand, the U.S. methodology has been criticized for using a complex mathematical algorithm to achieve equivalence of values across Sections. The criticisms have involved both the details of the mathematical procedure itself as well as the fact that few people were able to understand what was done, how, or what effects the linkage process may have had on the original values. On the other hand, not using such an algorithm has been criticized for substituting subjective judgements for science and mathematics. Needless to say, there is no one solution that will please everyone.

The methodology we will use attempts to minimize and simplify, as much as possible, the need for linkage at the end of the process by:

i. facilitating the equivalence of reference codes at the beginning of the process by selecting reference codes according to common criteria, evaluating those codes by Section representatives at a working group session, and discussing the results of those evaluations with all Sections;

ii. facilitating equivalence between the survey code evaluations of each Section by using a consistent approach to evaluating all codes (i.e. same Likert scale of 1-7, same provision of speciality-specific reference codes, same definitions of inputs, common education of evaluators, common evaluation method etc.); and,

iii. providing a link between specialities by ensuring that, wherever possible and appropriate, codes common to more than one OHIP speciality are evaluated by each of those relevant OHIP specialities.

The Commission is consulting with its experts and will make its final decisions on linking once it has seen and analyzed the data resulting from the reference and survey code evaluations.

Fitting in unevaluated codes *(see p. 53 of May 1998 report)*

Question: How will fitting be done and why hasn't this been decided yet?

Answer: First, it's worth bearing in mind that the codes to be evaluated account for approximately 90% of professional billings.

There are a number of accepted methods for fitting in unevaluated codes. The Commission continues to be of the view that it would not be practical or useful to choose a method before reviewing the data. The Commission is consulting with

its expert panel on what basis a particular fitting method should be chosen.

Redefinition *(see p. 16 of May 1998 report)*

Question: Some issues that the Commission has labelled as part of the redefinition process need to take place now, not later. Why is this not being done?

Answer: The Commission is very sensitive to the concerns of many that the OHIP Schedule is in need of significant redefinition. Some want codes to be removed, others want services to be added. Some believe the code descriptions are no longer adequate or applicable in some cases. Changes to the provision of medical services, such as recognizing the complexity of some physicians' patient profiles, may be more appropriately addressed in the Schedule. For a RBRVS to work, however, the existing Schedule, with all of its possible shortcomings, needs to be first expressed in relative value units. After that, the work of redefinition can begin. The Commission will recommend to both the Ministry of Health and the OMA that the work begin soon and be implemented as quickly as possible.