

Objective and Reproducible Model Assessment

Robert J. Isaacson

The ubiquitous orthodontic model has been part of orthodontics as long as we have been looking at orthodontic questions. Still we have no evidence based way of rating a model. Standard practice is to hold up or mount a set of models for visual inspection and offer an opinion presumably based on our individual concept of what constitutes ideal or perfection. We are biased by Angle's "old glory" concept, but we are aware of individual variation.

What is missing is an agreed upon gold standard individualized for this patient. Everyone knows of tooth size issues, problems that arise with upper premolar only extractions, etc. Still it is almost a given that no group of orthodontists will agree on exactly what the best outcome is—in fact the same orthodontist will not always agree with yesterday's decision upon subsequently repeating the assessment. It is common knowledge that model assessment is a problem for the ABO clinical examinations. In fact reliable model assessment is a need for any third party payer who must prioritize orthodontic treatment.

The second major problem is the need to pick points in order to measure models. The quick and common procedure was to have the evaluator simply pick up the model and make an estimate of the variations using a personal concept of where the ambiguous point was located. We have long understood that there is an envelope of error associated with point picking and when you are asked to identify a point on a round or ambiguous image, the envelope is potentially large. The outcome, of course, is that no two people or even the same person can pick exactly the same point twice in a row. Thus the parameter under measurement has considerable variation making model parameter scoring not reproducible.

These problems were recognized by the ABO and, to their credit, they developed a measuring tool that avoided the eyeball estimation technique commonly done. This was a constructive move toward objectifying the model assessment problem. Unfortunately, the tool still needed to have points selected and that retained all of the problems associated with point selection.

As with so many similar problems, an improved approach lay with the evolution of the digital world and creation of the digital model. A gold standard can be

set for each individual patient with a digital set up that is agreed upon by everyone before the treatment is decided upon. With this gold standard you can assess the difficulty of the treatment by measuring the difference between the assessment of the existing malocclusion as compared to an assessment of the gold standard proposed outcome on the setup. Later, the goodness of the treatment is assessed by comparing the finished (or long term retention) model to the gold standard set up.

We have been working with one software engineer devoted only to this question for several years. It is now possible to digitally superimpose arches and individual teeth and express differences in reproducible evidence based data. All that remains unknown is exactly which parameters should we be assessing and how much tolerance is acceptable in each one. Most importantly, a measurement is relative to something else and what is the reference is a question will become increasingly important with this major step forward in model assessment.

Imagine where this development will potentially impact your life. Obviously the ABO would be delighted to have a functioning tool with these properties. Indeed, with the growth of the number of diplomats and the commitment to periodic reexamination, such developments are mandatory. Past assessments of models have been limited to the external surface, but now many more parameters are accessible. Much research will be needed to establish the most important parameters to assess and the most important ones to treat. All third party payers will want to know how to prioritize their budget to approve the most useful and important treatments. The ABO difficulty Index DI is open for review. Is the difficulty for the treatment or is it which feature of the problem is most the most troublesome for the patient's long term outcome? It can be either one or any other depending on what you ask the system to give you.

Of especial importance is the long term nature of this data. Digital models can be stored indefinitely. Since the assessment is reproducible, the records years later can be compared with whatever parameters have evolved as most important from ongoing research. Any given parameter can be assessed at any later date and the result will always be the same

no matter who does the work. In essence, model assessment can be delegated to a lesser trained personnel with no difference in quality. The doctor is necessary to ask the question, set the methods and interpret the results. The models can be assessed without compromise by any trained personnel.

Everybody always gets the same result because the operator makes no judgments. The operator merely runs the system and everyone will get the same result irrespective of who does the measurement.

The plaster model enters the 21st century!