PERSPECTIVE

Artificial Intelligence (AI) as an Aid in Providing Clinical Care: A Personal Perspective

Paul Harris, OD, FOVDR, FACBO, FAAO, FNAP
Cordova, Tennessee USA

INTRODUCTION

It seems that everywhere we turn these days we encounter the two letters "AI" with promises that our lives will be made easier in ways we have yet to imagine. In many ways it turns out that these promises and AI itself are going through a cycle we have seen repeatedly in the medical field when something new is first discovered. Thomas Kuhn in his book, "The Structure of Scientific Revolutions[1]" laid out many of these steps. He was an historian of science and made explicit the steps by which science evolves and what must happen for new thoughts or ideas to take their place as the new "status quo" in a field. AI or Artificial Intelligence is the new kid on the block right now and we are just at the beginning phases, in many different arenas of thought and practice, of trying it on for size and seeing in what ways it can be helpful.

Correspondence regarding this article should be emailed to Paul Harris, OD at **Paul.HarrisOD@gmail.com**.
doi.org/10.31707/VDR2026.12.1.p28.

Harris P. Artificial Intelligence (AI) as an aid in providing clinical care: a personal perspective. Vision Dev & Rehab 2026; 12(1):28-43.

Keywords: Artificial Intelligence, expert systems, AI aided case analysis

In this piece I will share my thoughts on years of playing with AI, or at least a form of it, and where it may prove to be helpful in the long run, but also to warn us of becoming over reliant on this tool too soon in too many areas of clinical practice. By reading this perspective I hope you get the following:

1. AI didn't first come onto the scene just a few years ago. It's been around quite a while.[2]
2. AI is a much larger field than just what is in the news now: Large Language Models.
3. We all should be skeptical of that which is being attributed to AI.
4. When turned inwards on what we do every day, we need to be careful not to cede too much clinical judgement to AI.

AI holds promise in assisting us all in our daily lives, but we must always remain responsible for the actions we take and to not rely too much on the tools that will be emerging over the next few years.

Terminology and Categorization

Let's begin by defining some words and terms. These definitions come from a document published by Professor Christopher Manning of Stanford University's Human-Centered Artificial Intelligence in September of 2020.[3]

Intelligence: *Intelligence might be defined as the ability to learn and perform suitable techniques to solve problems and achieve goals, appropriate to the context in an uncertain, ever-varying world. A fully pre-programmed factory robot is flexible, accurate, and consistent but not intelligent.*

Artificial Intelligence: *is a term coined by emeritus Stanford Professor John McCarthy in 1955, was defined by him as "the science and engineering of making intelligent machines". Much research has humans program machines to behave in a clever way, like playing chess, but, today, we emphasize machines that can learn, at least somewhat like human beings do.*

And 68 years later, this definition from the OED (Oxford English Dictionary): *The capacity of computers or other machines to exhibit or simulate intelligent behavior; the field of study concerned with this. In later use also: software used to perform tasks or produce output previously thought to require human intelligence, esp. by using machine learning to extrapolate from large collections of data.*[4]

There are multiple kinds of AI, with the most prevalent now being what the public is calling "Large Language Models" (LLMs),[5] which the field of AI calls "Natural Language Processing". Examples of these LLMs include ChatGPT, Gemini, and Claude to name only three as of this writing. The main types of AI are Machine Learning, Neural Networks, Natural Language Processing, and Robotics.

There are three more categories of AI which currently seem to be referring to the whole field, but they could be used to describe the evolution of the fields in any of these four domains. These are Narrow AI, General AI, and Artificial Super AI. From Stanford University:

- ***Narrow AI:*** *Artificial intelligence can be built for very specific tasks, such as playing a game, keeping spam out of your inbox, helping you find a nearby restaurant with your smartphone, or even driving your car.*
- ***General AI:*** *With more resemblance to human capabilities, general artificial intelligence is a more advanced form that can involve visual and language processing, contextual understanding, and the ability to adapt to a range of tasks. It's considered to be far off in the future.*
- ***Artificial super AI:*** *Imagine a machine that's smarter than you — much smarter. Artificial superintelligence is still only a theory, but advances in nascent AI are raising interesting and troubling questions for the future of humanity.*

The Evolution of Artificial Intelligence

The next step in the evolution of AI came from exposure to expert systems. From the OED:

> **Expert Systems:** *are a piece of software programmed using artificial intelligence techniques. Such systems use databases of expert knowledge to offer advice or make decisions in such areas as medical diagnosis and trading on the stock exchange. An expert system is a* ***computer program*** *that uses artificial intelligence* ***(AI)*** *technologies to simulate the judgment and behavior of a human or an organization that has expertise and experience in a particular field.*

The concept of expert systems was developed in the 1970's by computer scientist Edward Feigenbaum, a computer science professor at Stanford University and founder of Stanford's Knowledge Systems Laboratory. The world was moving from data processing to "knowledge processing," Feigenbaum said in a 1988 manuscript.[6] That meant computers had the potential to do more than basic calculations and could solve complex problems thanks to new processor technology and varying computer architectures.

At this point in time there was no simple way to capture the knowledge of an expert. What was born was the field of "knowledge engineering". Knowledge engineering is a subset of artificial intelligence (AI) that tries to emulate the judgment and behavior of a human expert in a given field. Knowledge engineering is the technology behind the creation of expert systems to assist with issues related to their programmed field of knowledge.

Two seminal papers in this field are, "Catching Knowledge in Neural Nets" by Jorgensen, CC and Matheus, C. from AI Expert, December 1986 and "NETtalk: a Parallel Network That Learns to Read Aloud", from Johns Hopkins in 1986.[7]

Figure 1 shows a flow diagram for how NETtalk was to work. Text was input in a stream of letters and spaces in a manner that reminds one of modern day chyron used at the bottom of many news broadcasts. The neural network had to learn or to be taught. Text would be given and the sound

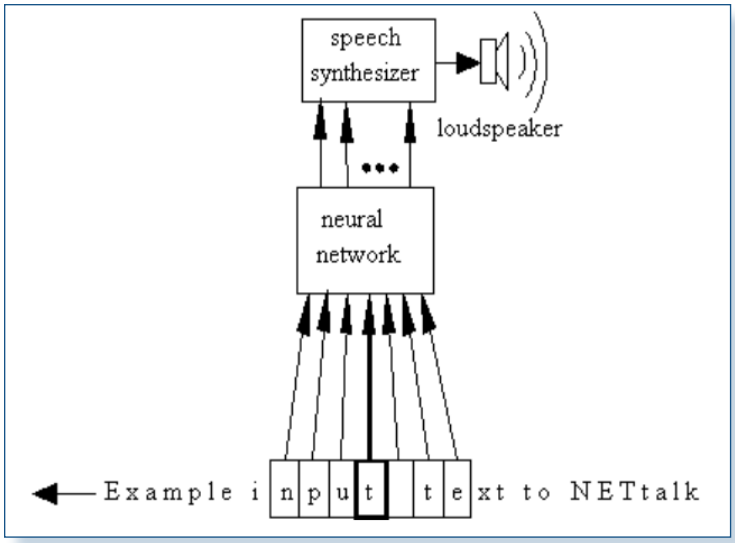


Figure 1. Plan for NET-talk.

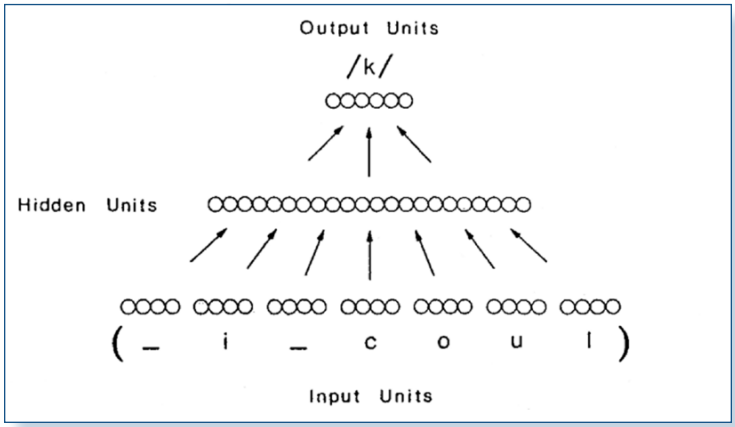


Figure 2. The Neural Network in NETtalk.

produced would be evaluated for correctness and if correct the weights in the connections between the input layer to the hidden layer and from the hidden layer to the output synthesizer would be reinforced. If a negative result was produced, the connections dominating would be weakened following the rules elucidated by Donald Hebb with his work on what happens at the synapse level in the brain.[8]

Figure 2 shows the neural network in NETtalk with the text to be decoded below "I coul", connected to the hidden units, which then were connected to the output unit at the top layer which, in this case, had the phoneme "k" for the first sound in the work word "could". If the system got it correct it was rewarded or reinforced by the programmer.

Those working on NETtalk did not know ahead of time what the tradeoffs would be between the

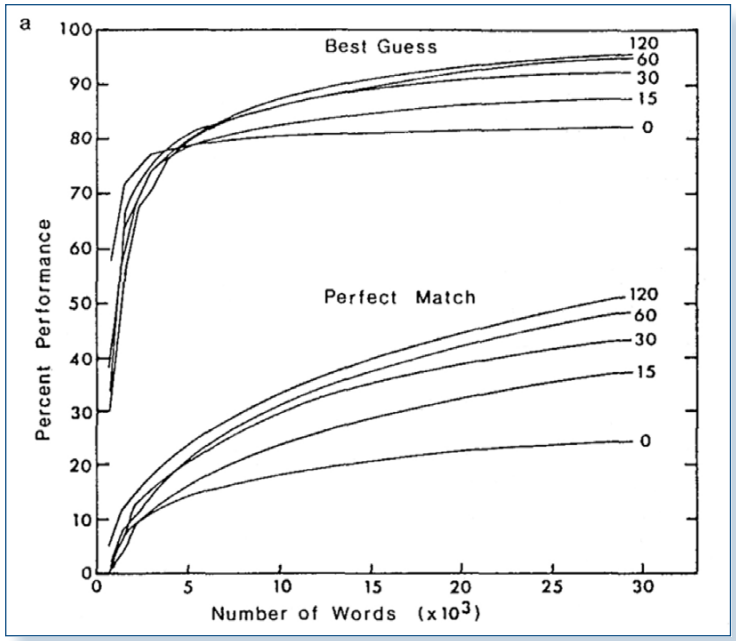


Figure 3. The learning curves and percentage of "best guesses" and "perfect matches" as a function of the number of words it was trained on and the number of hidden units built into the system.

number of hidden units, the number of teaching trials needed, and the degree of accuracy attained by the system. As a result of the work shown in figure 3, they learned that increasing the number of hidden units increases the learning capacity of the system, but there is an optimal number of hidden units, which in their model turned out to be 80. A benefit of this number of hidden units, and the authors were presaging what we have learned in clinical practice in working with patients who have suffered TBI's, is that the system is also quite robust in terms of being able to withstand damage to individual hidden units. As the number of hidden units damaged increases, the system still works well but slower at first and only after a significant number of hidden units are damaged does it begin to make errors.

Additionally, anticipating what we and allied fields do with our therapies, they did an experiment where they "damaged" a system that was already performing well. What they did was randomly alter the connections at each hidden unit up (stronger) or down (weaker) and then do "therapy" on the broken system. What they showed was that the system recovered very quickly with new learning. For those interested, their ultimate system comprises 203 input units with 26 output units and 18,629

different weights in the connections between input and hidden and hidden and output layers. With this they achieved a 95% level of accuracy on trained texts with an 80% accuracy on new texts. Remember, this was 1986!

At the time this triggered some thoughts:

- Once learned, information, either declarative or procedural, is highly resistant to hardware failure (i.e., loss of certain cells will not affect the performance of the network).
- Positive and negative aspects of information are automatically balanced as a network reorganizes to solve a problem (sounds like Piaget 's process of accommodation). This is describing the process of achieving structural balance. Positive connections represent agreement, support, or friendship. Negative connections represent conflict, opposition, or antagonism. A paraphrase from Bruce Wolff is applicable here. "The visual process is the only process in humans that allows for the comparison of likes and differences simultaneously."
- Abstraction of data occurs automatically as a byproduct of learning pattern recognition, which occurs in parallel and reconstructively.
- Hierarchical data structures can be conveniently represented as multiple layers of networks. Networks can exhibit properties reminiscent of adaptive biological learning and can select and generate their own pattern features from exposure to stimuli.
- Neural networks can capture patterns occurring both in time (for example, auditory information) and space (such as visual data) and can operate in discrete or continuous representation modes.

Many people take introductory computer programming courses, and they often begin by using the computer language BASIC to have the message, "Hello World" come up on the computer screen. This type of programming is called procedural, and it consists of step-by-step with conditionals such as "if-then" structures or "if-then-else" structures along with instructions to GOTO or JUMPTO other parts of the "in-line" code, depending on the language. See Appendix A for an example of this kind of code written by the author to help understand the flow of income from opening a VT practice. These were easy to write, and one could follow the path through the code to see that it all worked correctly or to find errors in the code, called bugs. But this kind of software couldn't represent the complex structures needed for the kind of neural networks that could make AI emerge. Thus, the need for non-procedural languages of which LISP and PROLOG are two examples.

The Application of Expert Systems

Expert systems have two major components: an inference engine and the database the engine can access to solve the problems presented to it. Inference engines have two different directions they can work on a problem called forward chaining and backward chaining. NOTE: When people solve problems there are nearly always elements of both forward and backward chaining occurring simultaneously and most likely in parallel.

Forward chaining: from the current state, generate possible futures, evaluate each of those possible futures, pick the ones that are the best and repeat till the problem is solved. In chess, one would have to generate all possible moves, then evaluate them and choose the best path. When there are so many possible paths, such as in chess, this limits how far ahead one can look. In optometry, we can think of the novice who is coming to a situation with little or no experience. They must run each test and collect that data and see where it leads them.

Bruce Wolff talked about what we do as experienced clinicians. He said that we simultaneously analyze and synthesize. He used to "analyze" for the taking in of the data from the probes we do and "synthesize" as integrating what has been gleaned from the results of the probes with what came before and the hypotheses about where one is heading.

The novice, who is still learning the exam and forward chaining their collection of data is analyzing with little simultaneous synthesis. They cannot have the piece of data they just collected be integrated with what came before and trigger shifts in the path through the data collection to prioritize the next most important probe. Only after all the data is collected can they begin synthesizing and begin understanding the person in front of them. A novice has no choice in this, as they don't have the database of experience to guide them.

Backward chaining: The person will make an educated guess that the solution to the problem posed is X or Y. Then the knowledge base is searched for specific evidence that either the hypothesis is supported or is not supported. If evidence supports the hypothesis, then the inference engine lets the user know that the problem has been solved. This reflects a level of expertise based on clinical experience. One of the things learned by the programmers of "Chess 3.5", a chess playing engine which came out of Northwestern,[9] was that chess Masters and Grand Masters use a type of backward chaining to get the chess pieces to advantageous squares by seeing patterns and moving pieces to get to better patterns. They did not have better memory or brute force thinking by generating more positions in their heads. They didn't look at "bad" moves or "mediocre" moves. They looked at patterns of pieces, based on their experience with the pieces.

In the medical world, the better diagnostician uses their cumulative experience to form hypotheses about a case and then they use these hypotheses to guide them to confirm or deny their diagnosis. A case in point from my early days as an extern supervisor is instructive. I was fully booked but a patient called with an emergency. I suggested that the patient come in right away and I let the extern do the initial "workup". The extern related that the patient's chief complaint was a dry eye, only one eye, which started while my patient was at the beach. I entered the room and knew with one glance that the patient had Bell's Palsy. I looked at him and asked him to smile and then to wrinkle his forehead to confirm the diagnosis. The student had no prior experience with the condition that would have been triggered by his forward chaining.

The Expert System MYCIN

In 1972 work began on MYCIN at Stanford University. The program was to be an aid to clinicians working in the field of infectious disease to help them make better diagnoses and derive more accurate treatments. The program could request further information concerning the patient, as well as suggest additional laboratory tests, to arrive at a probable diagnosis, after which it would recommend a course of treatment. If requested, MYCIN would explain the reasoning that led to its diagnosis and recommendation.

Using about 500 production rules, MYCIN operated at roughly the same level of competence as human specialists in blood infections and better than general practitioners. It might come back with a statement like: "It is 80% certain that the infectious agent responsible for your patient's current illness is a staphylococcus gram negative bacterium".

The work to develop this was done by a team, which consumed the amount of time a single person would have spent over 40 years accomplishing. It had built into it the ability to query the system to see what was relevant to it from the knowledge base. In a teaching environment, be it with an actual patient or in a simulated patient encounter, having access to the knowledge base and the rationale it used, could be helpful. It could keep us away from speculating on what goes on in the black boxes of both the patient as well as our own thinking.

Demystifying the Magic of Clinical Care

Many of us have externs who spend time with us. There are some predictable patterns which emerge over the course of their time with us. Early on, they observe and we seem to move so fast from question to question and from probe to probe. Our patterns may seem predictable for a few steps and then we deviate from what they think we are doing and they don't get the why. They might even think there is a randomness to it, when in fact each movement is guided by the clinicians' problem-solving algorithm, which the clinician might

not even be able to put into words. To the extern watching, what made an impression were the rules they had learned to follow that were being violated. Because they couldn't yet see the connections that got the experienced clinician from point A to point Z, it all seemed like magic to them.

There is no magic. Rather, there is clinical intuition. That intuition is based on simultaneous bidirectionality. Backward chaining emerges from pattern recognition, which occurs most often below the threshold level of the person. That constrains the person in communicating it both to themselves and others. Douglas Hofstadter is an American cognitive and computer scientist whose research includes artificial intelligence. He coined the term "sealing off", which is apt. This refers to a barrier which comes into play once a repeated pattern and larger understanding emerge. In many ways it is an "emergent", something new, which though being a product of all those things which feed into it, has become something new, more than the sum of all those parts.

Observations of performance, when seen by the expert with a large knowledge base, lead them to the most highly probable understanding in a short period through pattern matching. This looks like magic. If this could be made explicit, what would emerge is a definitive textbook of behavioral optometry, along with an interactive teaching and consultation tool for the profession. It would be a noteworthy contribution to apply AI systems that unlock the inference engine and knowledge base of behavioral optometry. This should enable more people to benefit from full scope optometric care.

Caveats in Case Analysis Aided by AI

In the summer of 1987, I attended the Joint Conference on Theoretical and Clinical Optometry (JCTCO) cosponsored by OEP and Pacific College of Optometry. Bob Yolton, OD, PhD, learned of my shared interest in the area of AI and Expert Systems and gave me a copy of the computer program "1st Class". 1st Class is an inductive expert system development tool written by William Hapgood and supplied by 1st-Class Expert Systems Inc.[11] To use this program, you put data into a database and next to that data you put the action you took or suggested. So, the expert puts in a series of "factors". A factor is anything that the expert thinks "might" be relevant to the derivation of a particular set of actions. Factors may be either word factors or number factors. Word factors are a finite list of things such as hair color. The possible values that could be taken would then be input under the heading of "hair color", such as brown, blonde, and black.

The other type of factor is numerical. Although the use of the term numerical might seem to imply that clear concise mathematical relationships must exist between the members of the factor list this is not implied in any way. This type of classification system is used for anything that is represented as a number. 1st-Class seems to handle most of these scales as simply ordinal rather than equal interval. It simply looks for threshold values above and below which can determine a difference that exists relevant to the action decision taken.

Within our ranks there have been many conversations around the scales we use, such as the prism scale for measuring phorias or the prism vergence ranges at distance and near and what the differences are between measures done with Risley prisms in a phoropter versus measures done with prism bars in "free space". It is beyond the scope of this paper to go into all critical points made by people over the years on this topic but only recently, a series of comments by Robin Lewis, OD, helped to bring much of the difficulty here into focus.

We were talking about pulling together data to turn AI loose on to see how it might help us. With computerized EHR systems, the thought that this might be easy were soon dashed by our musing. Rob made several salient points. He said that you couldn't directly compare the numbers from subject A to subject B. At first, I recoiled from this and fought for the position it seems the majority of eye care professionals operate from, which does take these numbers to be relatable from person to person and from time to time within the same person.

Then we recalled Robert "Bob" Sanet, OD, and Gary Etting, OD, speaking of their work with elite athletes. In general, their findings taken during the exam in an office with this population are often mediocre. Examples would be base-in and base-out breaks in the 12-14 range when the expected might be 21 or 22. However when these elite athletes engaged in their sport, their optometric findings met or exceeded normative values or the expected while under pressure.

We have all had the patient who is a poor observer on whom we get a base-out break point of 36 at distance without a blur point and without any comments by them on a size or distance change. Is that measurement of 36 PD actually better than the elite athlete with a break of 14 PD? Rob suggested that what we were dealing with may be just noticeable differences (JND's). It may be possible that the elite athlete has fine JND's throughout that range while the poor observer has very wide JND's. Based on their personal "response scale" the elite athlete has many more steps within their range that are significant than the patient with the 36 break point. This may correspond to finer spatial judgements and finer control of their movements and their images of achievement than the poor observer. Thus, in this direct comparison the 36 may not be "better" than the 14, though it is certainly a larger measurement.

So back to 1st Class and playing with clinical data. You must first seed the database with data and the action taken by the expert from that data. You do not need to give any of the rationale from the expert, just the action recommended or taken by the expert. 1st-Class will generate a rule from the data given. The most relevant method for us is the "optimize" method. The optimize method creates 'optimized' (compacted) decision trees by choosing the right factors in the right sequence.

1st Class could only look at one thing at a time, not an entire data set and come up with all the complex relationships at once. This was mostly a limitation of the computing power at the time. How to use this tool in optometry was a key question. It would have been wonderful to put in all the data collected on several hundred patients along with the actions taken and see what THE rule or rules were that it churned out. But the tool was simple back then and it could only address one relationship at a time.

When looking over cases with others one factor seemed to paramount to guiding clinical choices of the range of lenses available to address the patient's needs as well as what role might VT play in the overall treatment plan for a patient. On the lens side of this, the paraphrase from Bruce Wolff comes to mind, "I'm still searching for the person who could not benefit from the appropriate plus for near." If you are a member of the human race, you can benefit from plus at near. And then there are Robert Kraskin's alternative of care from "good", to "better" to "best", with all alternatives being given to all patients and they select the best alternative for them at that point in time. This is all to say that determining that appropriate plus for near along with the benefits to be derived in the short term from said plus as well as the benefits over time are part and parcel of the clinical tasks of every optometrist at every visit. This fits hand in glove with knowing and presenting VT within the context of how best to help the patient meet their unmet needs.

Table 1: The relationship between level of embeddedness and several clinical factors which guide us in patient care.

	Embedded	Unembedded
Prescribing	The more embedded the closer the final prescription must be to the refraction or subjective	The more unembedded, the wider the range of acceptable lenses through which the person can achieve good clarity
Benefits of plus for near	The more embedded the less benefit the patient derives from plus for near	The more unembedded, the more benefits the person gets from plus for near
Plus acceptance	The more embedded the less open to using plus for near is the patient	The more unembedded, the more they embrace the plus for near.
VT	The more embedded, the more VT will be needed to set the stage for change in just about any area of life	The more unembedded, the less need there is for VT to help the patient make large changes in life.

In my thinking at the time, determining the level of embeddedness seemed to be a place to start. Looking back now at the language used at the time (late 1980's to the early 1990's) and seeing where our use of the term has evolved to require some suspension of judgement. At the time, the quality of embeddedness fit into four categories: Highly Embedded, Embedded, Unembedded, and Highly Unembedded. Table 1 shows the language of the time along this continuum for four aspects of optometric care.

I started with a blank slate in 1st Class and began filling it with data that was made up for examples. Question writers for exams do this all the time. They put together a case to ask a question about. The newbie simply goes along looking for the data point that doesn't fit. But the experienced clinician looks at the case and knows right away that this is a made-up case that just doesn't occur in real life. It was interesting to see that what 1st Class returned from the first try with "made-up" data was worthless. This same lesson is surfacing repeatedly and more strongly now with many of the AI large language models in relation to the training data they have been turned loose on. When using simulated data, their responses are not helpful and they may even "hallucinate", give answers that don't exist.

The next step was to input 60 records of my own patient base. I put in the full Skeffington Analytical Sequence, which includes not only my refractive data but distance retinoscopy, my binocular balance (7), the best visual acuity "BVA" refraction (7a), phorias through the habitual and through the new 7a, distance base-out and base-in vergences,[9,10,11] the fused cross cylinder (14B), phoria through that (15B), the base-out and base-in vergences through the near control (16 & 17), and the PRA and NRA (21 & 20). Next to each dataset appeared in a column with my classification of one of the four categories of embeddedness.

1st Class generated a rule from the first 60 sample patients. To see how well it worked I gave it 10 new cases it had not "seen". It got 7 of the 10 in the same category I had them in and the other 3 were only one step off. I then "corrected" those three and had 1st Class make a new rule with the 70 subjects. On the next 10 test subjects we agreed 80% of the time. I repeated this several more times and once I reached about 100-110 subjects total it was agreeing with me 95% of the time. The rule it generated was 45 lines long. See appendix B for the entire rule. The first 22 lines of the rule are depicted in figure 4.

```
1: dv_bi_rec??
2: | <1.00:nv_bo_rec??
3: |    | <5.50:dv_bo_rec??
4: |    |    | <3.50:-------------------------hi_unembed
5: |    |    | >=3.50:nv_bi_rec??
6: |    |       | <5.50:------------------hi_unembed
7: |    |       >= 5.50:dv_bo_brk??
8: |    |           | <15.00:---------unembed
9: |    |           >= 15.00:---------embeded
10: |    >= 5.50:nv_bi_rec??
11: |       | <17.00:dv_bo_rec??
12: |       |    | <3.50:nv_bi_rec??
13: |       |    |    | <5.50:---------hi_unembed
14: |       |    |    >= 5.50:---------unembed
15: |       |    >= 3.50:dv_bo_brk??
16: |       |       | <15.00:nv_bi_rec??
17: |       |       |    | <5.50:----unembed
18: |       |       |    >= 5.50:----embeded
19: |       |       >= 15.00:--------embeded
20: |       >= 17.00:dv_bo_rec??
21: |           | <14.00:----------------embeded
22: |           >= 14.00:----------------hi_embeded
```

Figure 4. The first part of the rule generated by 1st Class in classifying patients on their degree of embeddedness.

It is informative to re-visit this after so many years. 1st Class looked at the data and shows that the first finding to look at is the distance base-in recovery. In the expected findings that are part of the Skeffington Analytical Sequence (SAS), the break is at 9 and the recovery is at 5. Here it states that if the finding less than 1 prism diopter, then stay in the upper part of the decision tree and next look at the near base-out recovery. If that recovery is below 5.5 prism diopters, then look at the distance base-out recovery. If that is less than 3.5 prism diopters (PD) the case is highly unembedded. So, three findings, all recoveries were enough with some cases to identify the quality of the case being highly unembedded.

Backing up one step, if the distance base-out recovery is greater than 3.5 PD then it is time to look at the fourth and final recovery finding, the near base-in recovery. If that is less than 5.5 PD, then again, the case is highly unembedded. In the

case where the near base in recovery is greater than or equal to 5.5 PD then the first non-recovery type of finding is looked at and it is the distance base-out break. If it is below 15 PD, then the case is unembedded. If it is greater than 15 PD, then the case is embedded, and so on through the findings.

Through the entire rule the only findings that are part of decision tree were either recoveries or break points. What was fed into 1st Class were all the findings that are part of the standard analytical done by the author, which included many other findings. None of those findings showed in the generated rule. Further investigation along these lines has remained dormant since that time. Having access to hundreds or thousands of sample cases seemed to be the main sticking point for moving forward. No time has been given to using the "rule" generated yet as a teaching tool and nor should it probably be. Though using it as fodder for study groups to have grist for the mill which might lead to new clinical insights might be helpful.

One caveat emerges whilst looking at this in the light of the understanding of our current day. We must watch for the dangers of positing what is happening inside the black box and putting that out as actually what is going on. We need to be careful when searching for teaching maxims that we don't create more problems along the way. The fault of us humans is thinking that once we find a "rule" that seems to work, is thinking that the world actually follows those rules. This does not mean that there is nothing to be learned by going down these paths, just that we should never expect to find all such rules in their final form.

Can AI Aid Lens Prescribing?

Back to the late 1980's now for the next step, looking at one aspect of lens prescribing. Specifically, the near add value on cases who selected to do VT from the beginning of their care. The same database of cases was used but this time the clinical result 1st Class was asked to investigate was the near plus add given. Here, 100 cases were not enough to get much that was meaningful or helpful. See Appendix C for the complete rule.

Here is the order of findings, with the ones that help the rule match what was done coming first:

- Associated cross cylinder at near, or the FCC (14b)
- The subjective (7), sometimes called the binocular balance, most plus after binocular balance to the first good 20/20 (6/6 or 1.0) visual acuity.
- Positive Relative Accommodation (20)
- Distance retinoscopy (4)
- Stress point retinoscopy
- Near phoria through the new distance refraction

This part of the rule generated some minus lenses for near, which had not been done clinically. Firstly, there were not enough sample cases here to cover the entire spectrum of cases. Secondly, powers had been entered non-uniformly with some absolute and some relative powers existing for in the same data field. The whole thing needed the data to be normalized. After going back to look at the database, it turns out these were just a few of these cases and it skewed the top part of the rule. However, if we look at the bottom portion of the rule in figure 5, things make a lot more sense.

```
>=1.13:subject7??
    IF<0.625:phoria15b??
    |    IF<3.50:stresspnt??
    |    |    IF<0.875:-----------------0_75
    |    |    >=0.875:-----------------1_00
    |    >=3.50:--------------------------1_00
    >=0.625:nra21??
        IF<-0.125:subject7??
        |    IF<0.875:stresspnt??
        |    |    IF<1.13:---------0_75
        |    |    >=1.13:---------1_25
        |     >=0.875:subject7??
        |        IF<1.13:---------1_00
        |         >=1.13:stresspnt??
        |            IF<1.00:----1_25
        |             >=1.00:----1_70
        >=-0.125:assxcyl14b??
            IF<2.63:-----------------1_50
             >=2.63:-----------------grt_thn_3
```

Figure 5. The lower part of the rule generated by 1st Class where it was determining the near plus add to be prescribed for cases enrolling into VT.

Here the list of probes seems to make more sense clinically, and things that were expected to be relevant indeed are shown in the rule.

- Associated cross cylinder at near, or the FCC (14b) being greater than +1.13, meaning +1.25 or higher.
- The subjective (7), or binocular balance of +0.62 as the branching value
- The phoria through the 14b or FCC, the 15b with 3.5 exophoria

In the entire rest of the rule the only other findings to show up as part of the rule include the negative relative accommodation (21) and the stress point retinoscopy.

As taught, we typically see the maximum plus stress point retinoscopy value being about 0.25 more plus than the FCC or 14b. Many of the findings shown in this rule are ones that increasingly in the profession seem to be omitted from assessments. The data simply is not being collected. In those cases where plus at near is given, they must be basing the amount of plus on some other criteria. Research based on AI assisted rules may help to shed light on which derivations of plus for near correlated with best performance and least amount of effort and energy, though a fairly large number of subjects would be required to generate this. One other key aspect that is not factored in at all is the idea that as clinicians evolve from novices to experts, they "harvest" more clinical insights from fewer probes and these "insights" carry more weight over time than the hard numbers generated from the probes themselves. We don't know, to what extent, AI will know what data was taken from what developmental level clinician nor how to state its suggestions at the development level that clinician will understand.

Where are We Now?

Attempts have been made by several sources to amass a database of clinical data from which the kinds of insights which might have advanced the field could be made. One major effort was begun by the late Harold Haynes, OD at Pacific University College of Optometry.[12] After one of the early JCTCO meetings held at the optometry school, he enlisted a cadre of attendees to send in clinical data on patients. Many of us did take the time to do so but there were questions, early on, from many of the attendees, most of which related to standardization, or lack thereof regarding how the data was being collected.

Questions like:

- Did we all conduct each probe the same way?
- Did we all use the same targets?
- Did we all use the same lighting?
- Did we all do the probes/tests in the same order?
- Did we all record the data in the same way?
- Were each of our just noticeable differences (JNDs) of what we were observing the same?

Consensus at the time was that the answer to these questions was a resounding "no". The group became distracted by the effort to replicate an idealized system of exact individual measurements rather than realizing that the clinician to clinician and even within clinician variabilities held some of the most significant pattern data that would be lost by making everyone attempt to do it the same. So much effort into doing it a certain way, that many were then stripped of key aspects of how they gained insights into how the person in front of them was interacting with the world. As many of the volunteers wrestled with these questions, many decided they couldn't or shouldn't fulfil their commitments to send in cases and the effort died. Few other similar initiatives got any further. Perhaps the shift in generating useful tools should be away from grouped data from people who work in disparate settings and use different targets and lighting, and instruction sets, to start first with a single practitioner. Over the years, many optometrists seem to hone their test taking sequence, and when observed by a novice they may look routine and the same every time, but even that perception may not be true.

One Such Variation — Speed of Movement of Prisms

An age old question asked in our circles has been, "How fast do you move the prisms during both the measurement of heterophoria or prism vergence ranges?" In 2008, as part of the CITT study they used 2 prism diopters per second.[13] In conjunction with the Conference on Clinical Vision Care (CCVC), held at Southern College of Optometry in association with OEP held in 2012 a study was conducted to answer this question.

A standard phoropter with a student who would sit as a patient was set up in the research space. A video camera was set up to record what was done but its view was isolated on the phoropter. The 15 highly experienced practitioners, with an average of 25+ years of practice, who also regularly took both phoria and prism vergence ranges were asked to perform their sequence of tests from directly after the "refractive" part of the exam was finished, until they would pull away the phoropter from the patient. Lenses were left in through which the "patient" could see 20/20 binocularly at distance. The practitioners were told only that we wanted to get their sequence of testing and their instruction sets recorded. Nothing was mentioned about the speed of prism movement.

The data obtained was extremely enlightening. Figure 6 shows the average speed of prism movement per eye per second for each of the 15 practitioners. The standard deviation for each appears in brackets both above and below the average. Though we end up with a nice average of 1.5 prism diopters per eye per second, with a standard deviation of 0.75 PD, a few things popped out which could not be seen in this graph (right top).

Many of the practitioners repeated findings here and when they did, they often varied the speed of the movement of the prisms from one trial to the next. Seeing the following chart will help bring these variations to the fore.

In Figure 7 we have the speed for each trial taken by each practitioner with the speed for their first trial appearing in the left most column. Each practitioner is given their own unique shape and color, which you can see in the legend to the right of the graph. Variations on the first prism movement varied from just over 0.5 PD per second per eye (PSPE) for doctor 12 to just under 3.5 PD PSPE for doctor 11.

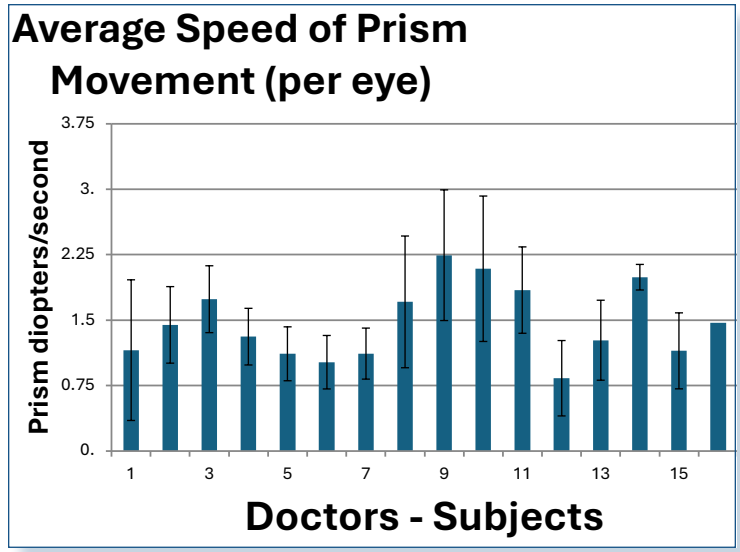


Figure 6 shows the average speed of movement of the prisms for each of the 15 subjects in prism diopters per second per eye.

What is so interesting is for example how doctor 11 varies the speed of prism from trial to trial. They began at just under 3.5 PD PSPE then went to 2.1, 1.9, 1.3 and then 0.75 PD PSPE for the next 4 trials. Doctor 15 had the least variation

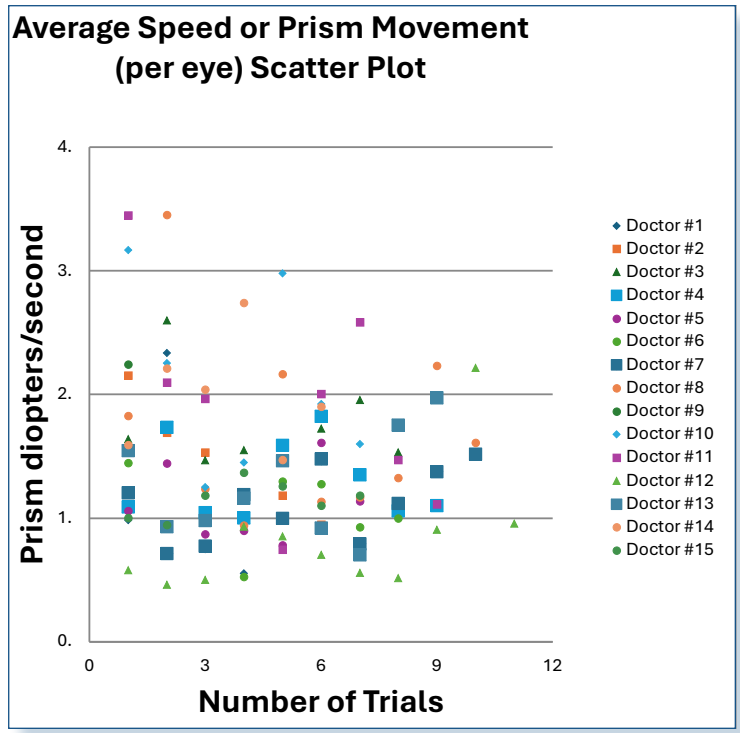


Figure 7 shows the speeds for each of the individual binocular prism probes done by our experienced panel of 15 doctors.

of all. Doctor 3 repeated the greatest number of prism probes, thus having 11 separate measures.

It was exactly these types of variations, both between doctors who do what we do, and even within the same doctor who changed speed a great deal from one probe to another with the same "patient" that continued to leave doubt that much could come from attempting to build a single big database of clinical numbers and have meaningful conclusions be drawn. Results like these should bring to the fore skepticism towards building large databases of just numbers from our clinical findings.

The Arrival of Large Language Models (LLMs)

The first LLM to go public was ChatGPT on November 30, 2022. It was released as a research preview and began as a chatbot. At the time of this writing the following are the most prominent examples of these.

- **OpenAI:** GPT-5, , GPT-4, GPT-o4-mini
- **Google:** Gemini, LaMDA, PaLM, BERT
- **Meta:** Llama (including Llama 3)
- **Anthropic:** Claude (including Claude 4.1)
- **Mistral AI:** Mistral models (e.g., Mistral 7B)
- **xAI:** Grok (including Grok 5)
- **DeepSeek:** DeepSeek-R1, DeepSeek-V3
- **Alibaba:** Qwen (including Qwen 3)
- **EleutherAI:** GPT-NeoX
- **BigScience:** BLOOM
- **Databricks:** Dolly
- **Cerebras:** Cerebras-GPT
- **Technology Innovation Institute:** Falcon (e.g., Falcon-180B)
- **MiniCPM:** MiniCPM4

Each seems to have its specialty in what it may be helpful for. The world seems inundated with everyone and everything incorporating AI. If you don't have it, you are broadcasting that you are not with it and many are including some kind of AI to check that box for one reason or another.

While preparing for a presentation in the early days of COVD on placebos, I came across information on cycles of discovery. I stated, "In the first phase, the world reacts positively because it is new, and it is great. The feeling in the air is often that a magic bullet has been found and of course, there are no side effects." It feels like most of the time since that November 2022 introduction of ChatGPT that we have been in this positive frenzy. But then comes the second phase which I described as, "The honeymoon is over." We begin to take the time and ask if the shiny new is better than what came before. In the third phase, we begin to fundamentally question whether things are better than before and in the end many things go back to the prior status quo, and the new-fangled thing finds its uses, but they are far more limited than imagined during phase 1. There are many things which could be cited which demonstrate that in some respects we are in phase 1 still, but there are many reports where using AI is creating more work and more problems than saving time, effort and energy. We will continue to see different aspects of the adoption of AI run through these phases.

Now many of us are using chatbots which are helpful most of the time and are mostly connected to some of the Large Language Models (LLMs). We type a prompt or give a voice command or prompt and get something back. Most of the time the stakes are low, so that if what we get is off or is a made-up response, sometimes called an "AI hallucination", it doesn't negatively affect our life too much. This is all one form of the use of AI, which is ubiquitous right now and which will most likely continue to increase rather than subside. Each time we use it though, we need to be cognizant of the "truthiness" of the return and to not take what it "says" as the gospel.

In the clinical aspects of our days there will be AI instances that help but again, we need to be vigilant with monitoring responses when it might directly affect our patient care. Setting up a chatbot on your office website to help with setting up an appointment might be a perfect thing, but it could lead to making appointments at odd hours or overlapping types of appointments that shouldn't be overlapped, etc. Extensive testing to try to break such a system should be done well before it is rolled out on your site.

Using a chatbot to interpret or represent a series Frequently Asked Questions (FAQ's), which you might have written or licensed to have on your site or to point your site to, could be fraught with giving poor medical advice which you may be held responsible for. This is an area I would be extremely careful of right now and for the foreseeable future.

Where AI should have its largest effect is with some of the data we collect during exam which is in confined universes where numbers rule the day. An example of this should be in visual field testing. Here the companies could have massive databases of prior visual field results and the clinical actions taken by the optometrist over time on tens to hundreds of thousands of patients. Using AI to detect patterns in the 50+ points tested to know which points are the "canaries in the coal mine" to show the earliest changes and send red flags to the clinician should be a daily occurrence. And I'm not talking about marking the whole field or one or two numbers on the report as RED but giving rationale and probabilities of developing certain types of problems within a given amount of time.

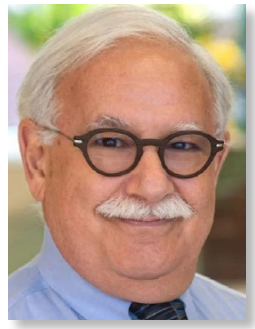
Using this kind of data should also massively reduce the time in testing in areas like the visual field. We should not have to do full thresholding at 50+ points on a retest or return test. Humphrey's STATPAC was a step in this direction, but it wasn't used to save any steps in testing. AI could review the patient's history and set up a protocol for that day, going right at the highly predictive areas of progression and then picking a few points in the rest of the area to confirm that things are still excellent or that a problem found before is still there to the same degree in both depth and area as before. Should a practitioner want new baselines, these AI driven custom protocols could be overridden periodically to "reset" the whole case for the clinician.

CONCLUSION

Joel Barker in his book, "Paradigms, The Business of Discovering the Future",[14] wrote, "If you want to be one of the first into the new territory, you cannot wait for large amounts of evidence. In fact, you must do exactly the opposite. If you want to be early, you must trust your intuition, you must trust your nonrational judgment and take the plunge; make the leap of faith to the new paradigm." Each of us needs to decide what level of faith to put into the AI we choose to use in our lives, and this will most likely shift greatly over time. How to not be fooled by an AI generated mass event like the, "War of the Worlds"[15] radio broadcast from 1938, which may look and sound far closer to reality than that Orson Wells, Mercury Theater broadcast? In our clinical lives we need to be ever vigilant that these tools are indeed just that, tools to help rather than mislead.

REFERENCES

1. ASIN: B00HMVLURM Publisher: The University of Chicago Press Publication date: January 1, 1994
2. Davenport, T, Kalakota, R., "The potential for artificial intelligence in healthcare", Future Healthc j. 2019; **https://tinyurl.com/3he2xwmx**
3. HAI Stanford University Human-Centered Artificial Intelligence Definitions – by Christopher Manning, September 2020.
4. Citation pulled 8/25/2025 – last OED revision 2023 for "Artificial Intelligence" **https://tinyurl.com/2evvdb7z**
5. Edmonds SA, Edmonds CT, Bolner M. Using large language models in neuroscience-optometry rehabilitation programs. Vision Dev & Rehab 2025;11(3):128-31.
6. **https://tinyurl.com/rzkaf9jb**
7. **https://tinyurl.com/mrxbya4j**
8. **https://tinyurl.com/5fjdjkbe**
9. **https://tinyurl.com/ycypenra**
10. **https://tinyurl.com/4sypdcux**
11. **https://tinyurl.com/2n9attrh**
12. **https://tinyurl.com/3nbw9x39**
13. **https://tinyurl.com/43rvwd93**
14. ASIN: 0887306470. Publisher: HarperBusiness. Publication date:May 26, 1993 ISBN-10: 9780887306471 ISBN-13: 978-0887306471
15. **https://tinyurl.com/2w6t966b**



AUTHOR BIOGRAPHY:
Paul Harris, OD, FOVDR, FACBO, FAAO, FNAP
Cordova, Tennessee USA

Dr. Paul Alan Harris is a distinguished optometrist and educator with a career spanning over four decades in behavioral optometry, clinical practice, and optometric education. He earned his Doctor of Optometry degree in 1979 from the State University of New York, State College of Optometry. Throughout his career, Dr. Harris has made significant contributions to both academic and clinical aspects of optometry, most notably during his tenure at the Southern College of Optometry.

APPENDIX A

```
11 PRINT "HOW MANY WEEKS WOULD YOU LIKE TO RUN?"
15 INPUT A
20 DIM B(A),C(A),D(A),F(A),J(A),K(A)
25 PRINT "HOW MANY PATIENTS WILL ENROLL INITIALLY IN THE VT PROGRAM?"
30 INPUT B(0)
35 PRINT "WHAT IS THE AVERAGE LENGTH OF VISION THERAPY IN WEEKS?"
40 INPUT Q
45 PRINT "WILL THE NUMBER OF NEW PATIENTS INCREASE?"
50 INPUT A$
55 IF A$="YES" THEN 80
60 LET S=1000.0
65 GO TO 100
80 PRINT "BY HOW MANY PATIENTS WILL THE NUMBER INCREASE?"
85 INPUT R
90 PRINT "AND AT WHAT INTERVAL IN WEEKS?"
95 INPUT S
96 PRINT "WHAT IS THE MAXIMUM NUMBER OF NEW VT PATIENTS?"
97 INPUT Z
100 PRINT "WHAT ARE THE PROFESSIONAL FEES PER VT VISIT?"
105 INPUT E
110 PRINT "WHAT ARE THE WEEKLY OVERHEAD EXPENSES?"
115 INPUT G
180 PRINT "  WEEK NEWPTS  TPTS   PROFEE/WK   TFEEYTD   OVERHEADYTD
200 FOR N=1 TO A STEP 1
231 FOR L%=1 TO 20 STEP 1
233 IF N=L%*S THEN 900
235 NEXT L%
240 B(N)=B(N-1.0)
260 IF N>=Q THEN 800
600 C(N)=C(N-1.0)+B(N)
610 GO TO 850
800 C(N)=C(N-1.0)+B(N)-B(N-Q)
850 D(N)=C(N)*E
860 F(N)=F(N-1.0)+D(N)
870 J(N)=G*N
880 K(N)=F(N)-J(N)
885 FOR$ ="  ###    #.##  ##.#   #####.##  ######.##  ######.##  ####
895 PRINT USING FOR$;N;B(N);C(N);D(N);F(N);J(N);K(N)
898 NEXT N
900 IF B(N-1)>=Z THEN 240
905 B(N)=B(N-1.0)+R
910 IF S>=Q THEN 800
920 IF S<Q THEN 600
999 END
```

APPENDIX B

This is the embeddedness rule from the 110 subjects done in 1987 by the author.

```
1: dv_bi_rec??
2: | <1.00:nv_bo_rec??
3: |     | <5.50:dv_bo_rec??
4: |     |     | <3.50:-------------------------hi_unemb
5: |     |     | >=3.50:nv_bi_rec??
6: |     |         | <5.50:------------------hi_unembed
7: |     |         >= 5.50:dv_bo_brk??
8: |     |             | <15.00:---------unembed
9: |     |             >= 15.00:---------embeded
10: |     >= 5.50:nv_bi_rec??
11: |         | <17.00:dv_bo_rec??
12: |         |     | <3.50:nv_bi_rec??
13: |         |     |     | <5.50:---------hi_unembed
14: |         |     |     >= 5.50:---------unembed
15: |         |     >= 3.50:dv_bo_brk??
16: |         |         | <15.00:nv_bi_rec??
17: |         |         |     | <5.50:----unembed
18: |         |         |     >= 5.50:----embeded
19: |         |         >= 15.00:--------embeded
20: |         >= 17.00:dv_bo_rec??
21: |             | <14.00:----------------embeded
22: |             >= 14.00:----------------hi_embeded
23: >= 1.00:nv_bo_rec??
24:     | <6.50:dv_bo_rec??
25:     |     | <6.00:-------------------------unembed
26:     |     >= 6.00:nv_bi_rec??
27:     |         | <5.50:------------------unembed
28:     |         >= 5.50:dv_bo_brk??
29:     |             | <15.00:---------embeded
30:     |             >= 15.00:---------hi_embeded
31:     >= 6.50:dv_bi_rec??
32:         | <7.00:dv_bo_rec??
33:         |     | <5.00:------------------embeded
34:         |     >= 5.00:dv_bo_brk??
35:         |         | <15.00:nv_bi_rec??
36:         |         |     | <4.00:----embeded
37:         |         |     >= 4.00:----hi_embeded
38:         |         >= 15.00:------------embeded
39:         >= 7.00:dv_bo_rec??
40:             | <10.50:dv_bo_brk??
41:             |     | <15.00:--------hi_embeded
42:             |     >= 15.00:nv_bi_rec??
43:             |         | <11.50:--unembed
44:             |         >= 11.50:--hi_embeded
45:             >= 10.50:------------------hi_embeded
```

APPENDIX C

This is the complete rule for prescribing the near add on cases who elected to do VT.

```
 1: assxcyl14b??
 2: IF<1.13:subject7??
 3: |    IF<-0.435:assxcyl14b??
 4: |    |    IF<0.500:pra20??
 5: |    |    |    IF<1.63:retinosc4??
 6: |    |    |    |    IF<-2.38:--------- -3.25
 7: |    |    |    |    >=-2.38:stresspnt??
 8: |    |    |    |        IF<0.625:-- -0.50
 9: |    |    |    |         >=0.625:--- -0.25
10: |    |    |    >=1.63:phoria13b??
11: |    |    |        IF<-0.500:-------no_rx
12: |    |    |        >=-0.500:-------0_25
13: |    |     >=0.500:retinosc4??
14: |    |        IF<-0.875:---------------0_75
15: |    |         >=-0.875:stresspnt??
16: |    |             IF<0.685:-------0_50
17: |    |             >=0.685:-------1_25
18: |    >=-0.435:assxcyl14b??
19: |        IF<0.875:stresspnt??
20: |        |    IF<0.685:retinosc4??
21: |        |    |    IF<0.375:--------0_75
22: |        |    |    >=0.375:--------0_50
23: |        |    >=0.685:----------------0_75
24: |         >=0.875:#embeded??
25: |            =hi_unembed:-------------1_25
26: |            =unembed:----------------0_75
27: |            =embeded:----------------1_00
28: |            =hi_embeded:-------------no-data
29: >=1.13:subject7??
30:     IF<0.625:phoria15b??
31:     |    IF<3.50:stresspnt??
32:     |    |    IF<0.875:-----------------0_75
33:     |    |    >=0.875:------------------1_00
34:     |    >=3.50:--------------------------1_00
35:     >=0.625:nra21??
36:         IF<-0.125:subject7??
37:         |    IF<0.875:stresspnt??
38:         |    |    IF<1.13:---------0_75
39:         |    |    >=1.13:---------1_25
40:         |     >=0.875:subject7??
41:         |        IF<1.13:---------1_00
42:         |         >=1.13:stresspnt??
43:         |            IF<1.00:----1_25
44:         |             >=1.00:----1_70
45:         >=-0.125:assxcyl14b??
46:             IF<2.63:-----------------1_50
47:              >=2.63:-----------------grt_thn_3
```